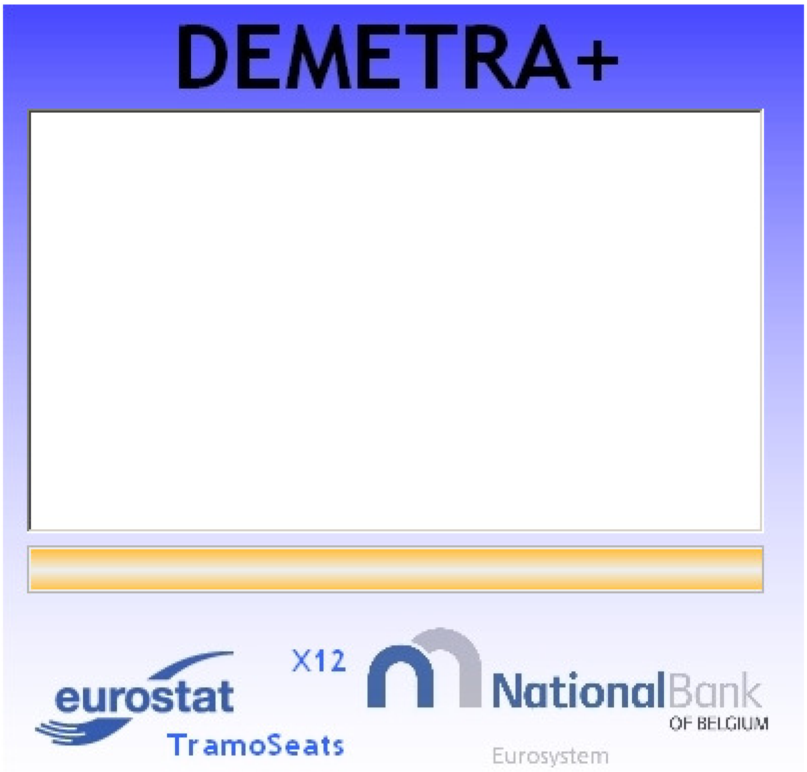
# Demetra+

# User Manual




Sylwia Grudkowska

National Bank of Poland

March, 2011

Acknowledgements:

I would like to thank all the members of the Steering Group on seasonal adjustment for their useful comments and helpful suggestions on various drafts of this document.

Thanks are due to Dominique Ladiray (INSEE), Jean Palate (National Bank of Belgium), Anna Ciammola (ISTAT), Faiz Alsuhail (Statistics Finland), Dario Buono (EUROSTAT), Joerg Meier (Bundesbank), Michael Richter (Bundesbank) and Kevin Moore (ONS) for their valuable support and contributions to the preparation this manual.

# Contents

# Introduction

Seasonal adjustment (SA) is an important step of the official statistics business architecture and harmonisation of practices. Since the 1990s, the Eurostat has been playing a role in the promotion, development and maintenance of an open source software solution for seasonal adjustment in line with established best practices. In 2008, European Statistical System (ESS) guidelines on seasonal adjustment have been endorsed by the CMFB and the SPC as a framework for seasonal adjustment of PEEIs and other ESS and ESCB economic indicators. ESS guidelines focus on two most commonly used seasonal adjustment methods – TramoSeats[1] and X-12-ARIMA[2] and present useful practical recommendations.

The aim of creating new seasonal adjustment software Demetra+ was to supply a flexible software solution which covers the recommendation of ESS guidelines in this area. Demetra+ was developed by the National Bank of Belgium. The application seasonally adjusts large-scale sets of time series and provides user-friendly tools for checking the quality of the SA results. Demetra+ includes two seasonal adjustment methods: X-12-ARIMA and TramoSeats.

TramoSeats is a model-based seasonal adjustment method developed by Victor Gomez and Agustin Maravall (Bank of Spain). It consists of two linked programs: Tramo and Seats. Tramo ("Time Series Regression with ARIMA Noise, Missing Observations, and Outliers") performs estimation, forecasting, and interpolation of regression models with missing observations and ARIMA errors, in the presence of possibly several types of outliers. Seats ("Signal Extraction in ARIMA Time Series") performs an ARIMA-based decomposition of an observed time series into unobserved components.

X-12-ARIMA is a seasonal adjustment program developed by the US Census Bureau. It includes all the capabilities of the X-11[3] program, which estimates trend and seasonal component using moving averages. X-12-ARIMA offers useful enhancements including: extend the time series with forecasts and backcasts from ARIMA models prior to seasonal adjustment, adjustment for effects estimated with user-defined regressors, additional seasonal and trend filter options, alternative seasonal-trend-irregular decomposition, additional diagnostic of the quality and stability of the adjustments, extensive time series modelling and model selection capabilities for linear regression models with ARIMA errors.

This manual aims to introduce the user to the main features of the Demetra+ software and to make the user able to take advantage of the powerful tools provided. This document presents an

---

1 TRAMO (Time Series Regression with ARIMA Noise, Missing Observations and Outliers) and SEATS (Signal Extraction in ARIMA Time Series) are programs supported by Bank of Spain, for more details see Gómez, V., and Maravall, A. (2001);, Caporello, G. and A. Maravall (2004).

2 X-12-ARIMA is supported by the US Bureau of Census, for more details see Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., Chen, B-C. (1998).

3 X-11 program was introduced in 1965. See SHISHKIN, J.,YOUNG, A. H., and MUSGRAVE, J. C. (1967).

overview of the software capabilities and of its main functionalities. Moreover, step by step descriptions how to solve some very basic tasks are included. It will give the possibility to reproduce results with user's own data. The guide shows the typical path to follow and illustrates the user-friendliness of the application. It is expected that the readers have already acquired background knowledge about concept of seasonal adjustment and are familiar with X-12-ARIMA and TramoSeats methods. For those readers interested in studying the seasonal adjustment methods in detail, bibliography is provided at the end of the manual.

It should be emphasized that Demetra+ makes use of X-12-ARIMA and TramoSeats algorithms, restricted with regard to their original implementations. For this reason, there are some differences between original programs and programs implemented in Demetra+. The aim was to develop the software which enables the comparison of the result from TramoSeats and X-12-ARIMA. For this reason, revision history and sliding spans analysis are available in Demetra+ both for TramoSeats and X-12-ARIMA. On the contrary, some functionalities implemented in original programs are missing in Demetra+ (e.g. using X-12-ARIMA under Demetra+ it is not possible: to choose different filters for each specific month, to do a preadjustment of the original series with "prior adjustment factors", to specify ARIMA model $(p,d,q)(P,D,Q)$ without some lags in the regular part[4]).

The User's Manual is divided into five parts.

Chapter 1 presents the general features of the software and installation requirements.

In Chapter 2 the application's menu is outlined. It is also shown how to visualize the data provided with the software and how to import new series from Excel.

Chapter 3 focuses on workspace menu and useful options offered by Demetra+.

Chapter 4 describes how to define the seasonal adjustment of a single series and many series. In this part the result of seasonal adjustment is discussed.

Some detailed aspects, like description of the tests and some technical issues, are described in the Annex.

4 For example, the user cannot specify the model (2,1,1)(0,1,1) without parameter AR(1).

# 1 Basic information

## 1.1 About Demetra+

The first release of Demetra+ contains Demetra+ itself (main graphical interface) and Excel add-ins: ColorAnalyser (a tool to search outliers in an Excel worksheet containing time series), Demetra+ XL (a seasonal adjustment tool in the Microsoft Excel environment, inspired by the Demetra+, which can be used for multiprocessing), XL Functions (Set of Demetra+ Excel functions). The add-ins are described in the documentation attached to the software.

The current version of Demetra+ uses the following core engines:

- TramoSeats dlls, dated 8/2009,
- X12 dll, used in Demetra 2.2.

The most important results (including the complete RegArima model) directly come from the core engines. All the diagnostics are computed outside the core engines (see below).

One of the strategic choices of Demetra+ is to provide common presentation/analysis tools for both TramoSeats and X12. So, the results can be more easily compared. This implies that many diagnostics, statistics, auxiliary results, etc. are computed outside the core engines. Demetra+ is of course highly influenced by the output of TramoSeats and of X12. Most analyses presented in Demetra+ are available in the core engines. However, the results with TramoSeats and X12 may slightly differ for a lot of reasons (different statistical/algorithmic choices, possible bugs). In any case the global messages on a seasonal adjustment are (nearly) always similar.

Amongst the most important tools implemented in Demetra+, the following functionalities should be mentioned:

- Likelihood (X12-like) / RegArima model (t-stat as in Tramo): RegArima model was re-computed in Demetra+ (X12, Tramo and "Stamp-like" solutions available in the framework),
- Residuals analysis (Tramo-like, but based on another set of diagnostics,
- Seasonality tests (X12-like),
- Spectral analysis (X12 definition),
- Sliding spans (X12),
- Revision history,
- Wiener-Kolmogorov analysis (Seats-like).

Such solution leads to more flexible software. New features are easy to add to the software without modifying the core engine. One of the key features of Demetra+ is the possibility to use the underlying algorithms through a rich application programming interface (API). This feature allows the integration of the routines in very different contexts as well as the making of new applications. The most important concepts (time series, seasonal adjustment...) developed to encapsulate the core engines are common to both algorithms. The code for making basic seasonal adjustment is straightforward. However, it is possible to use the API to solve very tricky problems. A minimalist example is provided in the Annex.

Amongst the peripheral services offered by Demetra+, the following ones should be stressed:

- Dynamic access to various "time series providers": Demetra+ provides modules to handle time series coming from different sources: Excel, databases (through ODBC), WEB services, files (text, TSW, USCB, xml...); the access is dynamic in the sense that time series are automatically "refreshed" by the software, which consults the providers to download new information. The model allows asynchronous treatment.
- Common xml formatting: the seasonal adjustment processing can be saved in xml files, which could be used to generate, for instance, WEB services around seasonal adjustment.

The software was designed to allow the adding of new modules without modifying the core application. The main features that can be enriched are listed below:

- Time series providers: the users could add their own modules to download series coming from other databases,
- Diagnostics on seasonal adjustment,
- Output of SA processing.

As mentioned above, the API could be used to generate completely independent applications, but also to create more easily extensions to the current application.

Demetra+ is compatible with Windows XP, Windows Vista and Windows 7. Although Demetra+ is a 32 bits application, it also works with 64 bits version of operating system.

## 1.2 Uninstall previous version of Demetra+

In order to remove any previously installed Demetra+ version, the user should take the following steps:

- Open the "Add/Remove Programs" function in the control panel,
- Uninstall Demetra+ if listed,
- Close the "Add/Remove Programs" function,
- Delete the Demetra+ home directory,
- Delete the program group/icons (if manually created).

## 1.3 Installing Demetra+

Execute the file "setup" and follow the instructions on the screen. Always take the default options, i.e. typical installation etc.

## 1.4 Running Demetra+

Start working with Demetra+, run Demetra+ via the newly installed Windows option under Programs, or start the Demetra.exe file directly from the Demetra sub-folder.

## 1.5 Closing Demetra+

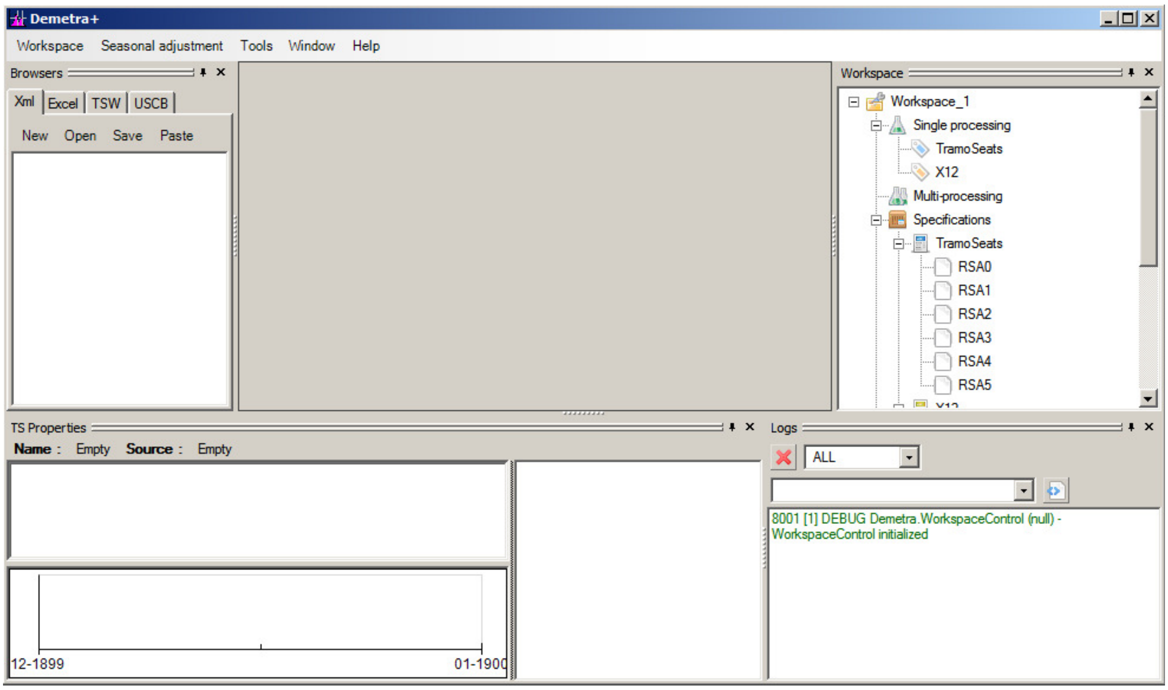
In order to close the application, the user can select File/Exit from the main menu (See Chapter 2). The other way is to click on the close box in the upper right-hand corner of the Demetra+ window,

If you created any unsaved work, Demetra+ will warn you and provide you with the opportunity to save it.

# 2 Main application's windows

## 2.1 Overview of the software

When the user launches the program he/she should see the Demetra+ window:




The key parts of the application are:

- the browsers panel (left panel), which presents the available time series,
- the workspace panel (right panel), which shows information used or generated by the software,
- a central blank zone that will contain actual analyses,
- two auxiliary panels at the bottom of the application; the left, one (TSProperties) contains the current time series (from the browsers' panel) and the right one (Logs) contains logging information.

Those areas will be described in next paragraphs.

Panels can be moved, resized, superposed and closed[5] depending on needs or preferences of the user. The presentation is saved between different sessions of Demetra+.

The application can contain multiple documents. Depending on the needs, the user can present them in different tabs taking the full space (default) or in floating windows (choose this one to follow different steps). The main menu item "Window->Floating/Tabbed..." gives access to that functionality[6].

5 Closed panels can be re-opened through the main menu commands: Workspace->View->...
6 See 3.5 for detailed description of this functionality.

Time series can be dragged and dropped between windows (next section presents how to do it). This function is omnipresent in Demetra+. It is the usual way to move information between different components. The objects that can be moved (time series, collections of time series...) can take different forms: nodes in trees, labels in lists, headers in tables, lines in charts...

When a drag and drop operation is initiated (which means that an object is indeed "moveable"), the cursor of the mouse changes to either a "no parking sign" or to a "+ sign", which indicates an acceptable drop zone.

Time series from Excel can easily be integrated in Demetra+. The users can create and import their own data sets. The series must be formatted in Excel as follows:

- True dates in the first copied column,
- Titles of the series in the corresponding cell of the first column,
- Empty top-left cell [A1],
- Empty cells in the data zone correspond to missing values (except at the beginning and at the end of the series).

That format corresponds with the format used by the Excel browser (which also requires the input zone to start at the beginning of the sheet [A1]). The exemplary file is presented below:

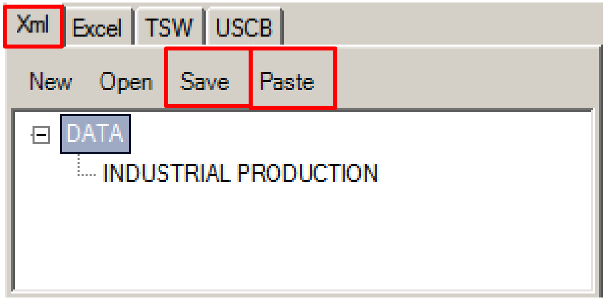
| | A | B | C | D |
|---|---|---|---|---|
| 1 | | Currency | M1 | M3 |
| 2 | 31-Dec-96 | 67865,98 | 140428,8 | 23563,88 |
| 3 | 31-Jan-97 | 63680,53 | 139384,1 | 22852,73 |
| 4 | 28-Feb-97 | 63625,74 | 141692,2 | 23513,54 |
| 5 | 31-Mar-97 | 65497,84 | 144931,6 | 24591,24 |
| 6 | 30-Apr-97 | 66635,33 | 148012,8 | 25873,43 |
| 7 | 31-May-97 | 69033,24 | 151700,6 | 25934,04 |
| 8 | 30-Jun-97 | 71672,27 | 154747,6 | 26835,17 |
| 9 | 31-Jul-97 | 74386,61 | 160454,2 | 27841,01 |
| 10 | 31-Aug-97 | 74328,59 | 162408,5 | 27907,99 |
| 11 | 30-Sep-97 | 74658,3 | 165037 | 27630,6 |
| 12 | 31-Oct-97 | 74854,52 | 170176,5 | 27663,32 |
| 13 | 30-Nov-97 | 75283,86 | 173413,4 | 27694,36 |
| 14 | 31-Dec-97 | 79239,77 | 179602,4 | 27255,87 |
| 15 | 31-Jan-98 | 73597,48 | 178239,7 | 26487,78 |
| 16 | 28-Feb-98 | 74417,05 | 180850,5 | 27403,85 |
| 17 | 31-Mar-98 | 75621,69 | 183236,3 | 27286,09 |
| 18 | 30-Apr-98 | 75681,04 | 185907,5 | 28834,44 |
| 19 | 31-May-98 | 78728,63 | 191080,2 | 28860,73 |

Time series are identified by their names. Information like data periodicity, starting and ending period can be derived from the first column.

After they have been copied in Excel, the data can be integrated in Demetra+ as follows:

- Select the Xml panel in the browsers,
- Paste the data (they appear in the tree). This option doesn’t work if some files were previously opened via Xml browsers. In this case, select the button **New** first and then **Paste**,

- Change the names of the series/collection in the tree if necessary (click twice on the item you would like to modify),
- Save the file (if need be).




Demetra+ is compatible with Excel 2003 and Excel 2007.

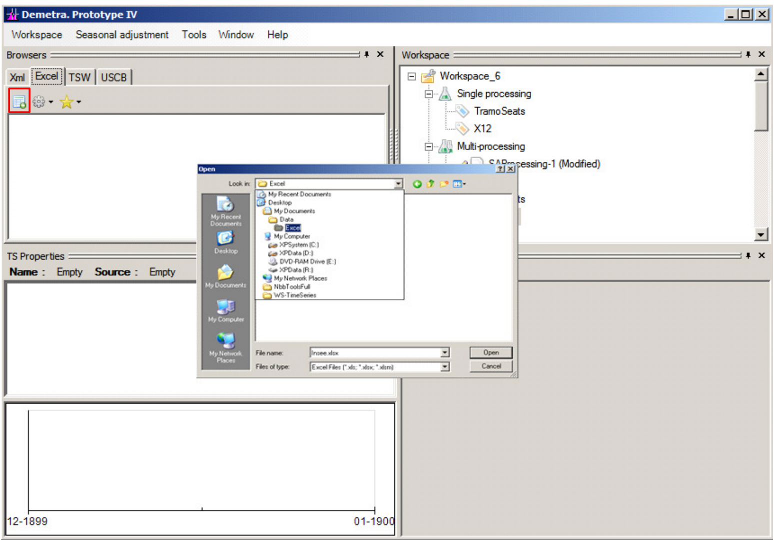
## 2.2 Browsers

The browsers' panel presents the series available in the software.

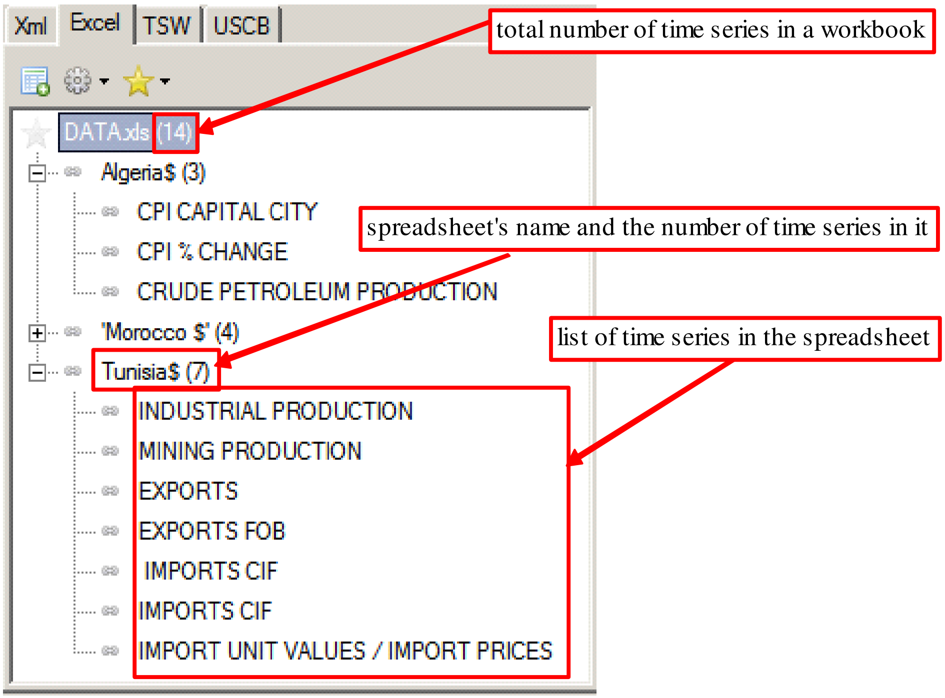
Different "time series providers" are considered: Xml (specific schema), Excel, TSW, USCB, Text and ODBC.

The installation procedure has copied several files in different formats in the subfolders of "My Documents\Data". The way how to open Excel workbooks is presented below. The procedure is similar for the other providers.

1. Click on the Excel tab of the browsers panel,
2. Click on the left button (see below),
3. Choose an Excel workbook (for instance "insee.xlsx") in the folder "My Documents\Data\Excel".



Final nodes of the trees represent time series and their parents represent collections of time series. Those nodes correspond with spreadsheets names. Different browsers show the data in trees that can be expanded by double-clicking their nodes (or single-clicking the "+/- signs"). The tree shows not only how the time series were organized in Excel's workbook, but also how many series are in whole workbook and in each particular spreadsheet.

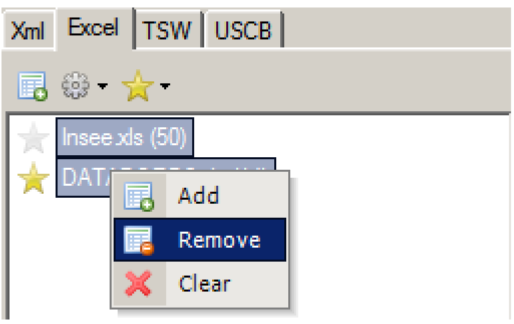



The right click on any time series name opens the pop-up menu, which contains the following commands: **Add**, **Remove**, **Clear**.

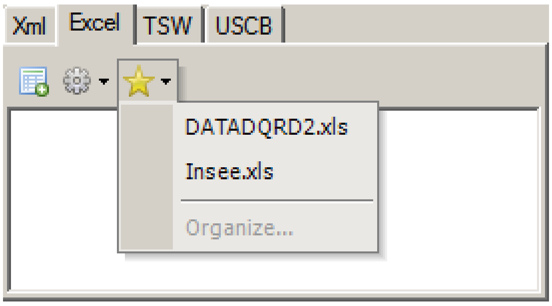
**Add** – opens new time series set form the Excel workbook,

**Remove** – removes all time series from the workbook. The button is active only if the name of the workbook is marked. It is not possible to remove all workbooks at the same time,

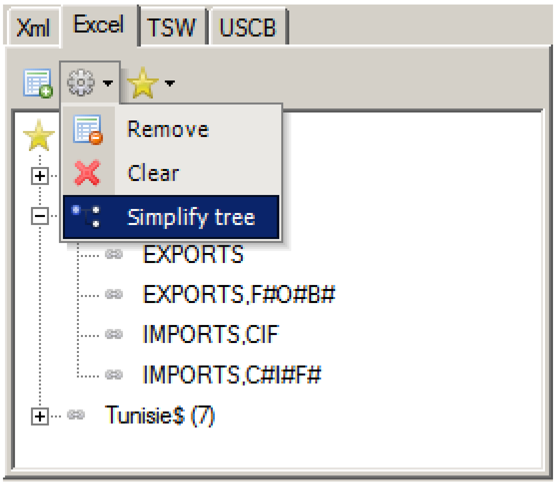
**Clear** – cleans the browsers' window.



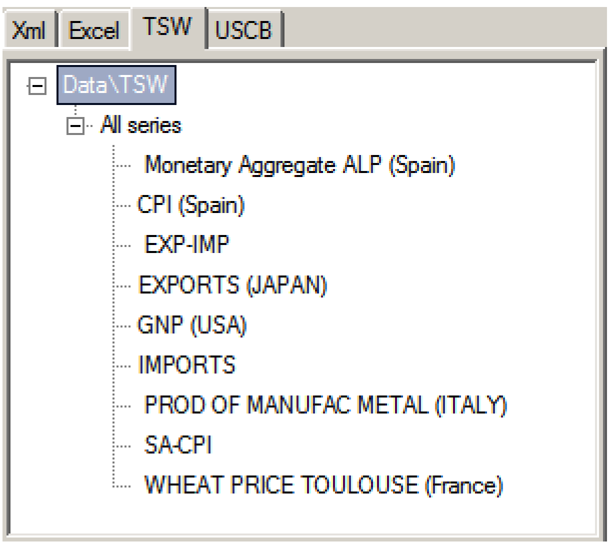
If the user wants to put the workbook into cache memory he should activate the star next to the Excel's workbook name. The list in the *Star* menu contains all workbooks, which are currently in the cash memory.




Using *Tool* icon (see below) one can remove marked item or clear the window. The **Simplify tree** option collapsed tree with opened branches.



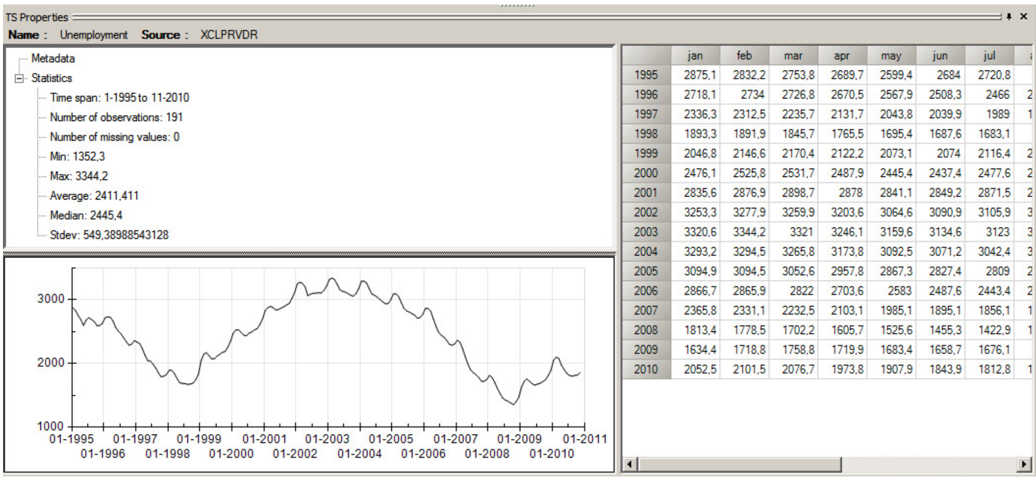

Demetra+ is able to read time series files written for TSW. The TSW folder can contain several levels of sub-folders with TSW files. They will appear in the tree navigator of the TSW provider. The series in a subfolder will be grouped in a collection called *All series*. The same idea was applied for USCB bookmark.



## 2.3 TS Properties

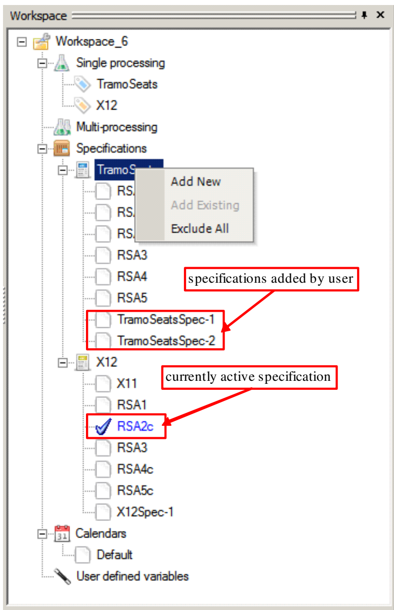
*TS Properties* window (contraction from Time Series Properties) can be used for examination the characteristics of individual series. This panel is strictly connected with the *Browsers*. The window is presented at the bottom part of the picture below.

*TS Properties* window presents the basics statistics, chart and time series data. The function is launched by single clicking on the time series name in the *Browsers* window. *TS Properties* provides also information about the name and source of the time series displayed in it.




## 2.4 Workspace

*Workspace panel* organizes all specifications as well as the processing and variables defined by the user. In the specification section some specifications are already defined. The user can add new specifications by choosing **Add New** from the pop-up menu (right click on the seasonal adjustment method's name). In *Workspace panel* the user can also define calendars and regression variables. The windows in which the user can define or change the seasonal adjustment parameters, calendars and regression variables will be described in Chapter three.



The right click on any existing name opens the pop-up menu, which contains the following commands: **Open**, **Exclude**, **Delete**, **Clone**, **Active**.

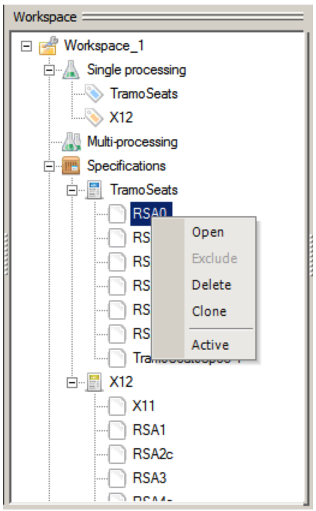
**Open** – opens the specification window with information about parameters. The user can't change them. The same result is achieved by double click on the specification's name,

**Exclude** – remove the specification marked. It works only for specifications defined by the user,

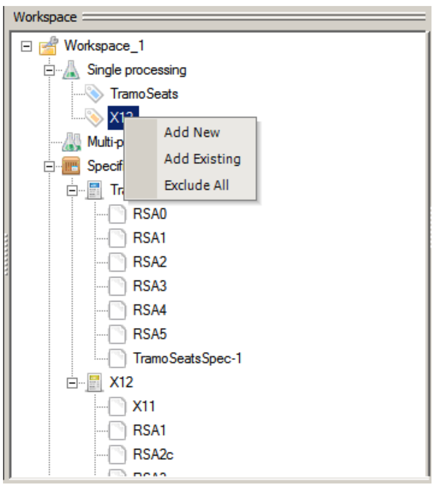
**Delete** – remove the specification marked. It works only for specifications defined by the user,

**Clone** – creates new specification, identical with the marked one. The parameters of the newly created specification can be edited by the user,

**Active** – activates the specification chosen. Time series will be seasonally adjusted using this specification.

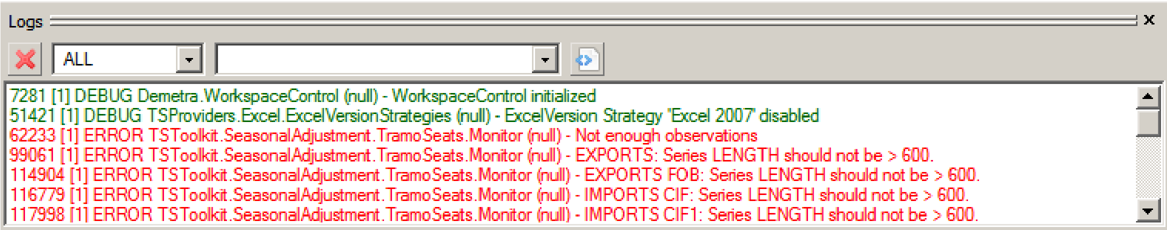



In a similar way the user can add new specification in single processing and multi-processing sections. This can be achieved by right-clicking on the seasonal adjustment method.

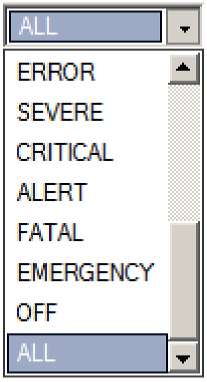


## 2.5 Log

*Log* window keeps information about all bugs, warnings and other events that took place during session.




The user can also display messages which belong to a chosen category (like ERROR, EMERGENCY – see the picture below).

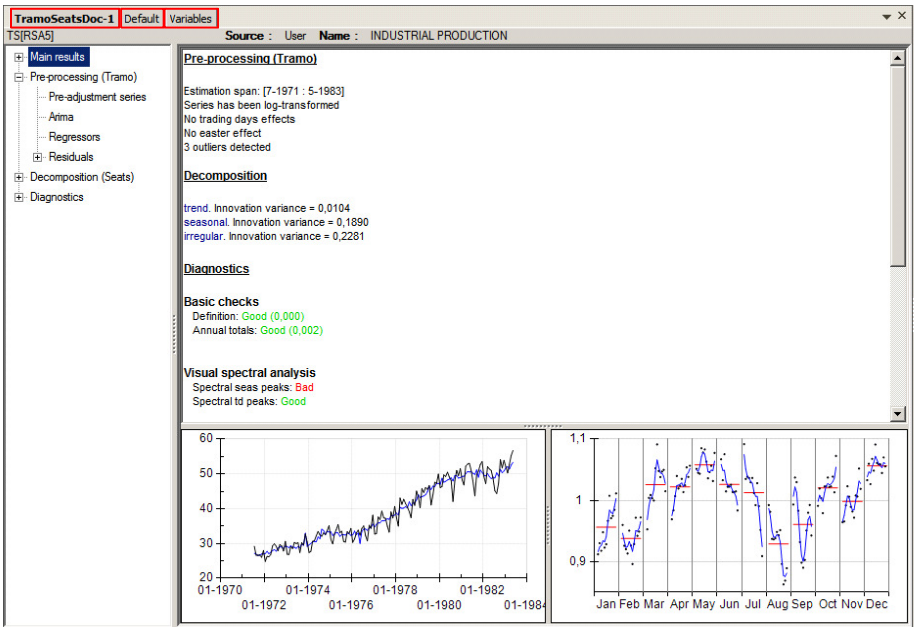



## 2.6 Results' panel

The black area in the middle of the window is the place where Demetra+ displays the various object windows that it creates. There could be displayed more than one window. Those windows will overlap each other with the foremost window being in focus or active. Only the active window has a darkened titlebar.

The example below shows the typical view of this panel. The right part of the panel presents navigation tree while on the left the actual results are displayed.

The user can execute several seasonal adjustments and define some regression variables. The results are displayed in consecutive bookmarks, which allow the user to switch them over. On the picture below it is shown that tree panels are opened - window containing seasonal adjustment results ("TramoSeatsDoc-1"), default calendar ("Default") and user defined variables ("Variables").


TramoSeatsDoc-1
Default
Variables
TS[RSA5]
Source : User Name : INDUSTRIAL PRODUCTION
Main results
Pre-processing (Tramo)
Pre-adjustment series
Arima
Regressors
Residuals
Decomposition (Seats)
Diagnostics
Pre-processing (Tramo)
Estimation span: [7-1971 : 5-1983]
Series has been log-transformed
No trading days effects
No easter effect
3 outliers detected
Decomposition
trend. Innovation variance = 0,0104
seasonal. Innovation variance = 0,1890
irregular. Innovation variance = 0,2281
Diagnostics
Basic checks
Definition: Good (0,000)
Annual totals: Good (0,002)
Visual spectral analysis
Spectral seas peaks: Bad
Spectral td peaks: Good
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

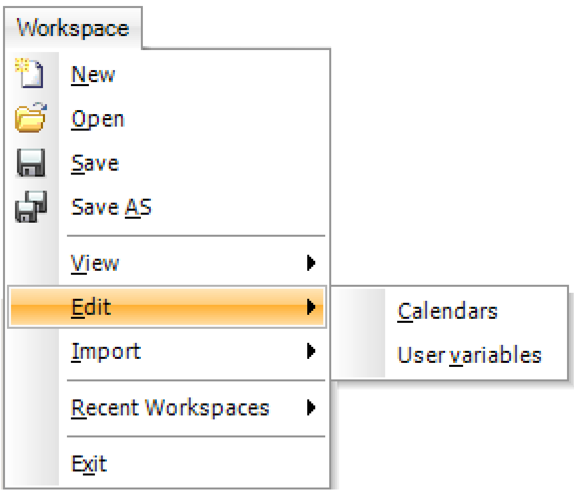
# 3 Application's Menu

The application's menu is situated at the very top of the main window. If the user moves the mouse's cursor to an entry in the main menu and click on the left mouse button, a drop-down menu will appear. Clicking on an entry in the drop-down menu selects the highlighted item. The functions available in application's menu are described in paragraphs below.

## 3.1 Workspace menu

Workspace menu offers the following functions:

- **New** - creates new Workspace displayed in the right panel,
- **Open** - opens an existing project in a new window,
- **Save** - save the project file named by the system (workspace_#number) that can be re-opened at a later point in time,
- **Save as** - save the project file named by the user that can be re-opened at a later point in time,
- **View** - activates or deactivates the panels chosen by user (Browsers, Workspace, Logs, TS Properties),
- **Edit** - allows defining countries' calendar and regression variables (this functionality is described further into this instruction),
- **Import** - allows importing countries' calendar and regression variables from Xml files,
- **Recent Workspaces** - opens workspace recently saved by user,
- **Exit** - closes an open project.



### 3.1.1 Calendars

This functionality is helpful for detecting and estimating the trading day effects. Trading day effects are those parts of the movements in the time series that are caused by the different number of the week in calendar months (or quarters, respectively). As with seasonal effect, it is desirable to estimate and remove trading day effects from the time series. Trading day effects arise as the number of occurrences of each day of the week in month (quarter) differs from year to year. The special case of the calendar effects is a leap year effect, which cause regular variation because of the extra day inserted into February every four years. These differences cause regular effects in some series. Both X-12-ARIMA and TramoSeats estimate trading day effects by adding regressors to the equation estimated in the pre-processing part (RegArima or Seats, respectively). These regressors are generated on calendar basis.

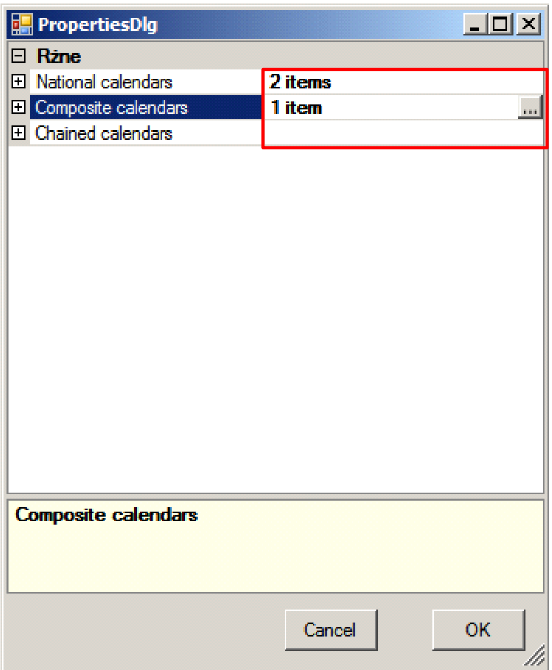
The calendars of Demetra+ simply correspond to the usual trading days contrasts variables, based on the Gregorian calendar, modified to take into account some specific holidays. Those holidays are handled as "Sundays" and the variables are properly adjusted to take into account long term mean effects.

Demetra+ considers three kinds of calendars:

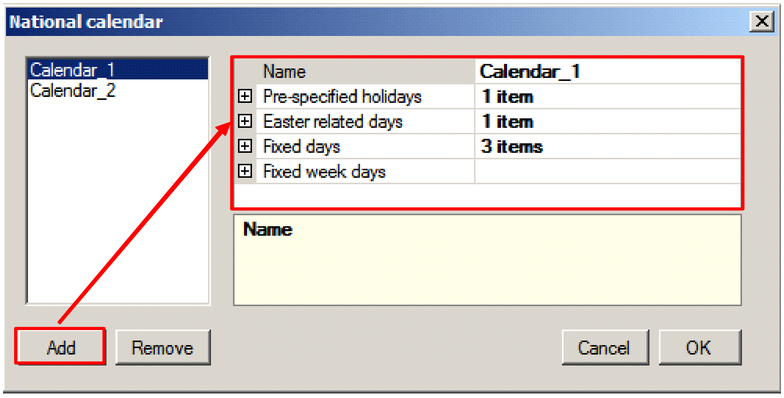
- National calendars, identified by specific days,
- Composite calendars, defined as weighted sum of other calendars,
- Chained calendars, defined by two other calendars and a break date.

The calendars can be defined recursively.

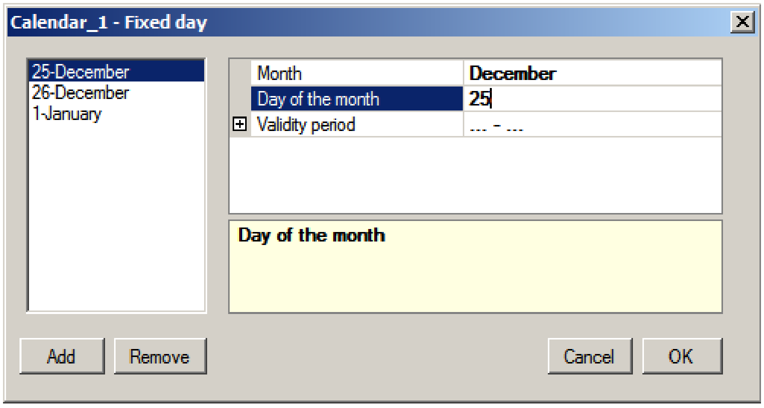
The dialog box allows defining all calendars described above. In the column on the right the number of calendars already defined is shown.



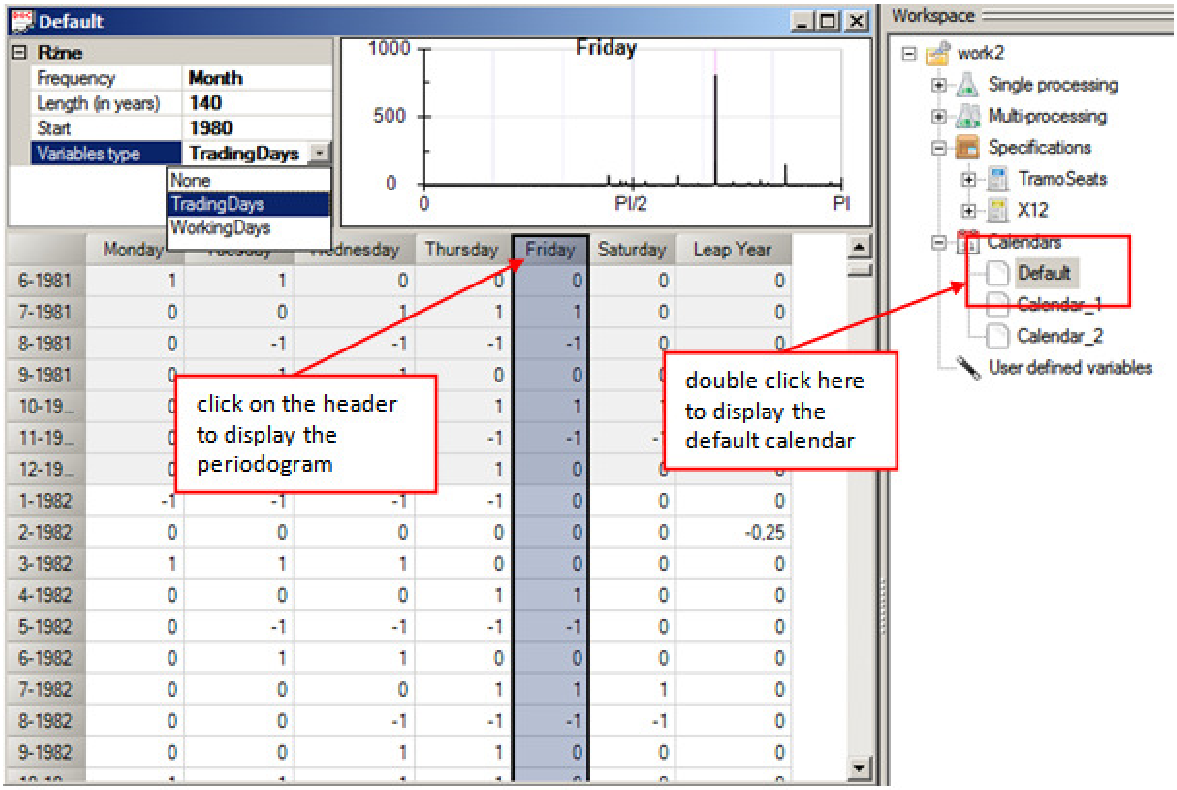
If the user chooses the option **National calendars** the following window is displayed. The user can define new calendar (**Add** button) or modify existing one. The list on the left contains all national calendars defined by user. In the panel on the right the user could specify the successive parameters.




In the example below it is shown how to define fixed holidays (choosing the month from the list and specifying the appropriate day of the month). If the validity period hasn't been specified, the regressor will be applied for all time series span.




The data generated by each calendar can be viewed by a double click on the corresponding item in the workspace tree.



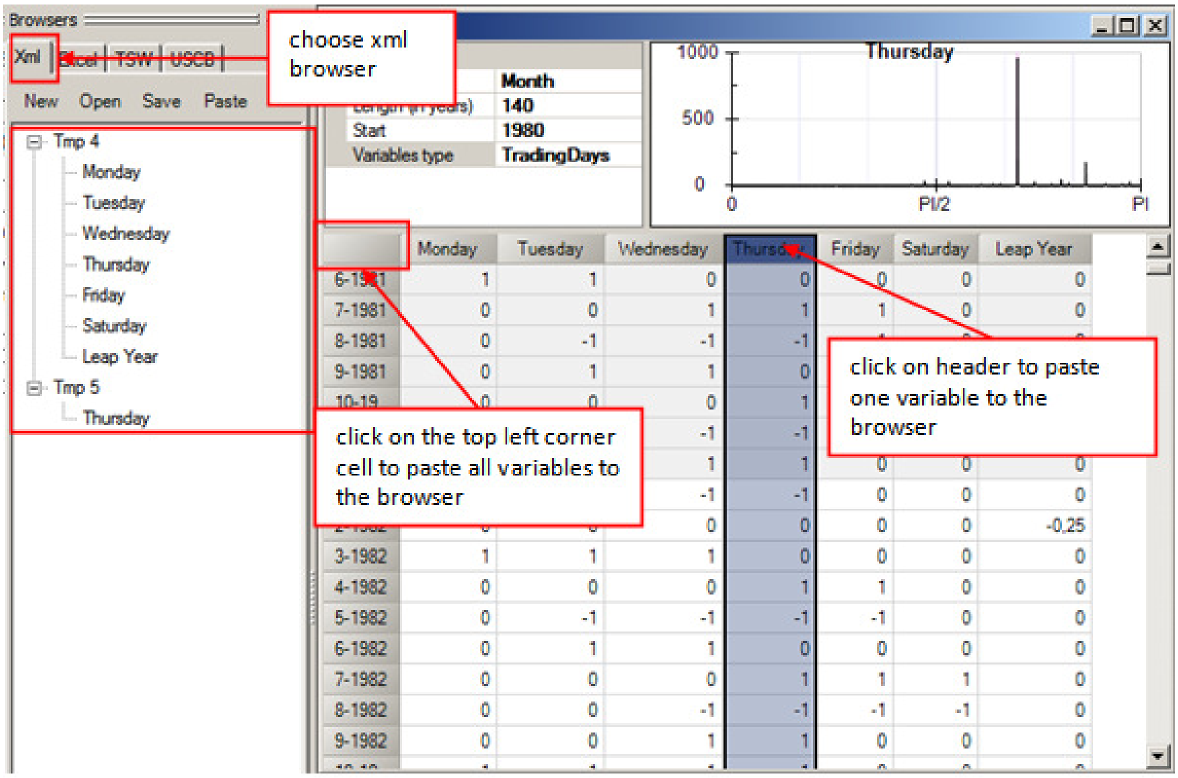

The regression variables can be inspected for any frequency (monthly, bi-monthly, quadric-monthly, quarterly, half-monthly, yearly) and any (reasonable) time span through that window; the periodogram of those series are displayed when a column is selected.

Demetra+ presents three different views:

- Trading Days - seven regression variables which correspond to the differences in economical activity between all days of the week and leap year effect,
- Working Days - two regression variables which correspond to the differences in economical activity between the working days (Monday to Friday) and non-working days (Saturday - Sunday) and the leap year effect,
- None - one regression variable which corresponds to the leap year effect.

This window should be used to analyse the data created by the calendar. Actually, Demetra+ enables the user to include/exclude the leap year effect from the seasonal adjustment model (see 4.1.3.1 and 4.2.2.1).

The series can be copied by drag and drop as it is shown in the picture below. The local menu can be used to copy and paste the series to other applications (e.g. Excel).




The calendars defined by the user are added to the *Workspace tree*. The user can display, edit or add new calendar by clicking on *Calendars* in *Workspace tree* and choosing appropriate option from the pop-up menu (for more details see: 2.4 *Workspace*).

### 3.1.2 User-defined regression variables

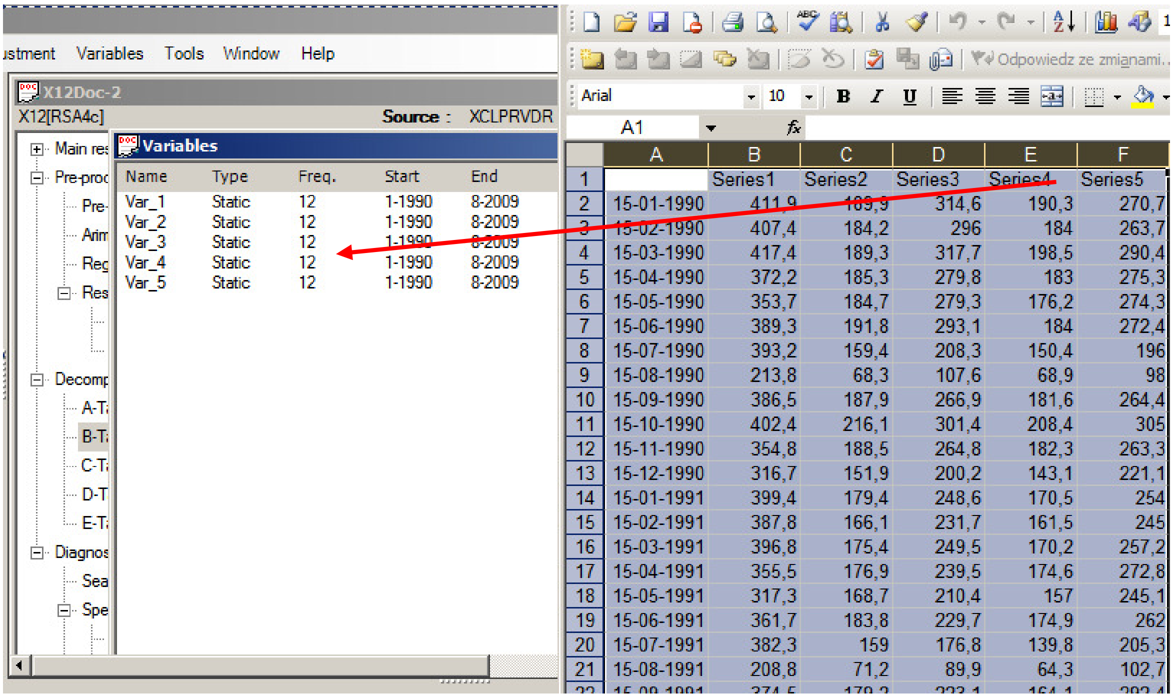
User-defined regression variables are simply time series identified by their name. Those names will be used in other parts of the software (regression) as identifier of the data.

Demetra+ considers two kinds of user-defined regression variables:

- Static variables, usually imported directly from external software (by drag and drop or copy - paste),
- Dynamic variables, coming from files opened with the browsers.

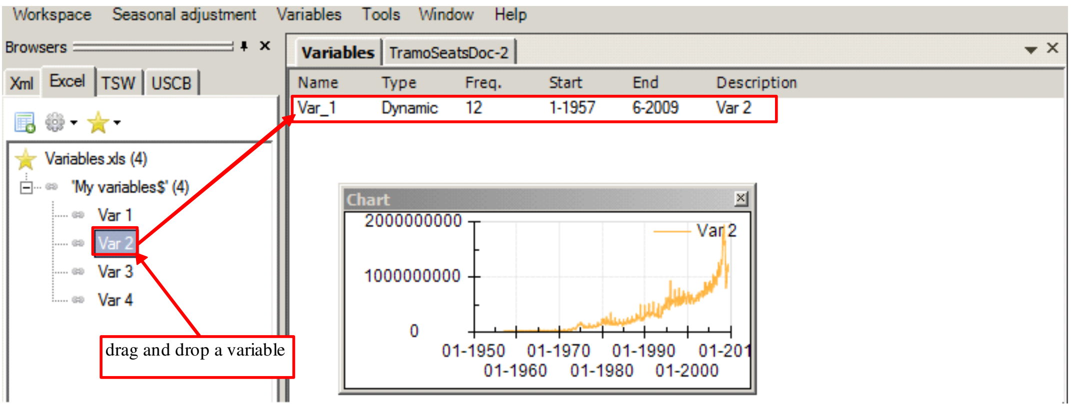
It should be emphasized that Demetra+ works on the assumption that a user-defined regressor is already in an appropriately centered form (i.e. the mean of a each user-defined regressor is subtracted from the regressor or means for each calendar period (month or quarter) are subtracted from each of the user-defined regressor).

Static variables imported directly from external software (for instance Excel) must be formatted as defined in the *Importing data from Excel* section. To import them, select User variables from Workspace menu (or double click item User defined variables in the Workspace tree) and by drag and drop time series from Excel or use the usual keys (*ctrl-c* and *ctrl-v*).




The figures of static variables cannot be changed. Currently, the only way to update static series consists in removing them from the list and to re-import them with the same names as previously.

Dynamic variables are imported by drag and drop series from a browser of the application.



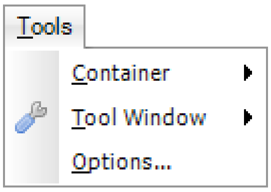

The names of the series can be changed by selecting a series and clicking once again when it has been selected. The selected series can be showed in a small chart window by a double click on regressor's name.

Dynamic variables are automatically updated each time the application is re-opened. Because of that it is a convenient solution for creating user-defined variables.

## 3.2 Tools menu

Tools menu is divided into tree parts:

- **Container** - tools for displaying data,
- **Tool Window** - charts and data transformation,
- **Options** - different windows, diagnostic and output options that can be set by user.




The contents of tool windows are automatically updated when:

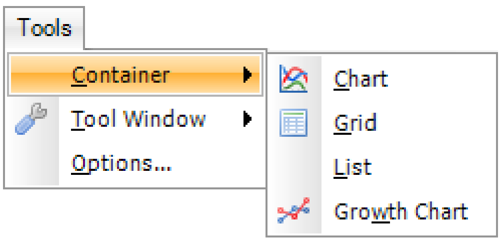
- A new series is selected, through a double click in the browsers panel or when a series is dropped in the left zone of the X12 window,
- The specification is changed, by means of the specification dialog box or when another specification, coming from the workspace, is dropped in the left zone of the X12 window.

Many other combinations are of course possible.

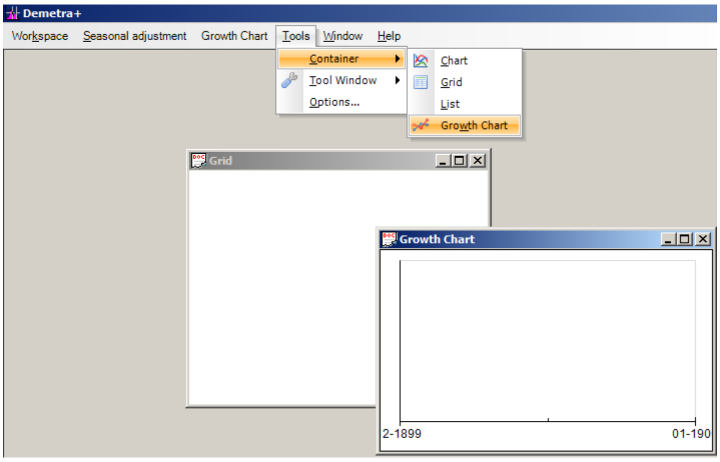
*Be advised that the current implementation is not able to detect recursive processing.. Such an attempt will generate a crash of Demetra+. The example of recursive processing is to select the series "D11" from X12 window and drop the series "D11 into the same X12 window from which "D11" was selected.*

### 3.2.1 Container

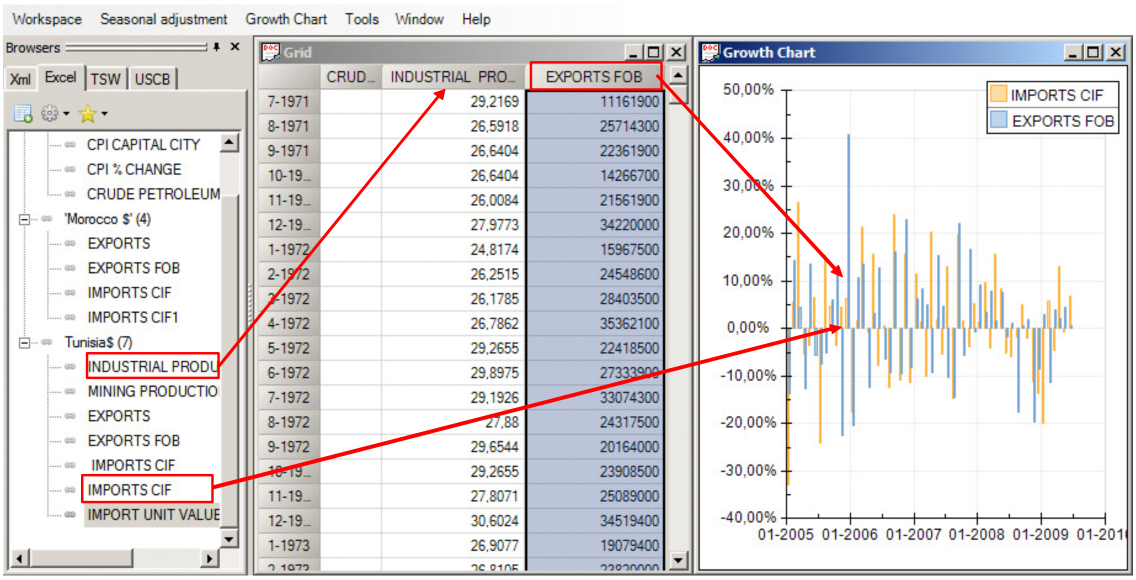
Container includes helpful tools to display the data. The following options are possible: **Chart**, **Grid**, **List**, Table or **Growth Chart**.




At first, the user should choose one or few containers from menu.

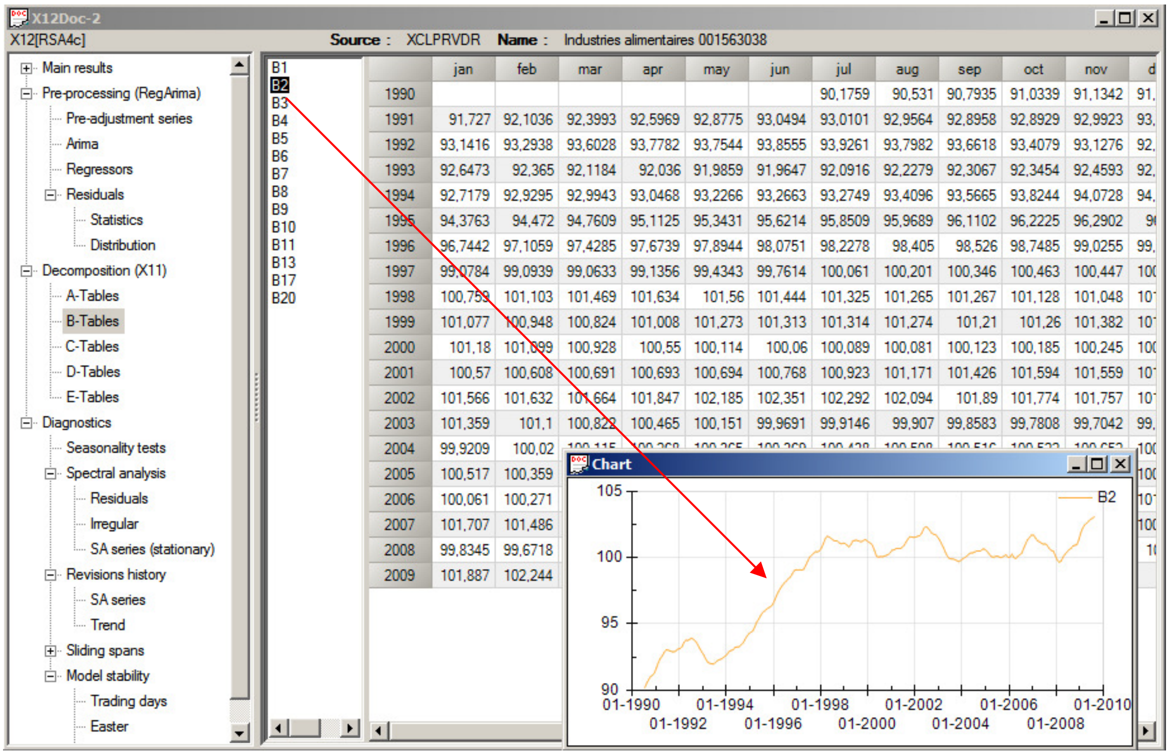



Then the user can take any series or group of series from one of the browsers and drop it in a container.



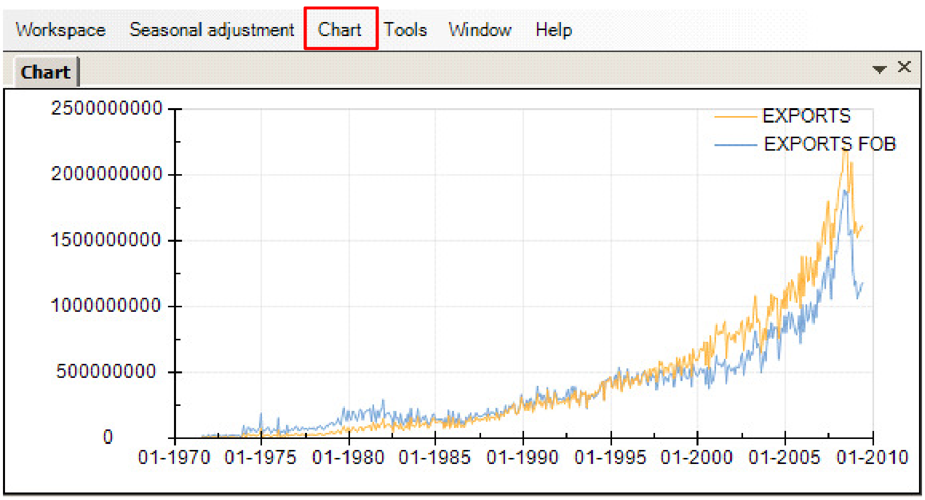

The group cannot be marked using **Ctrl** button from the keyboard. One can add the series to chart or grid, by dragging and dropping them one by one.

Different series, which appear in the results window (X12 or TramoSeats) can be dragged and dropped with the mouse to any other window of the Tools menu. It is also possible to drag and drop the results in the item chosen from container.

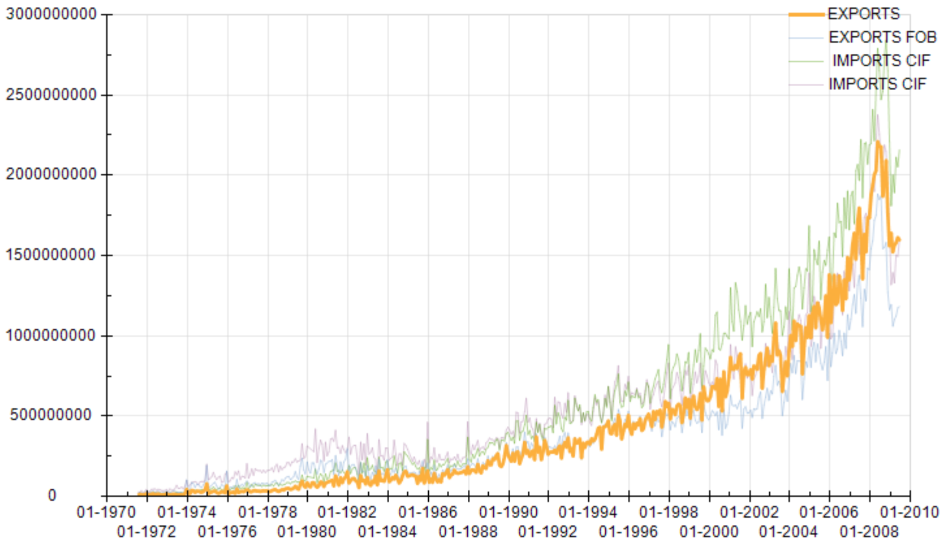



When a container is active, its name is added to the menu toolbar.

The chart (or growth chart) is automatically rescaled after adding new series. Also new item *Chart* (or *Growth Chart*, respectively) is added to menu toolbar.




Putting numerous time series into one chart could make it confusing. In this case the user can click on one series which is then displayed in bold.




The right-button menu offers many useful options. They content depends on the type of container. For example, for the growth chart the following options are available:

**Copy** – copies raw series and allows to paste it e.g. into excel. The function is active if the user clicks on the time series in the chart,

**Copy growth data** – copies m/m growth rates of the marked time series and allows to paste it e.g. into excel. The function is active if the user clicks on the time series in the chart,

**Remove** – removes time series from the chart. The function is active if the user clicks on the time series in the chart,

**Copy all** – copies all raw time series and allows to paste it e.g. into Excel,

**Copy all growth data** – copies m/m growth rates of the time series and allows to paste it e.g. into Excel,

**Remove all** – removes all time series from the chart,

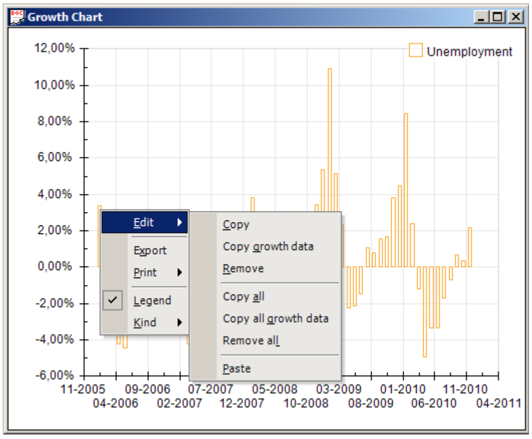
**Paste** – pasties time series previously marked,

**Export** – settings for export the chart, the option for chart can be copy to clipboard and save to file is also available,

**Print** – allows printing the graph and setting the print preview and printing page setup options,

**Legend** – add/removes legend from the chart,

**Kind** – displays *m/m* or *y/y* growth rates for all time series in the chart.

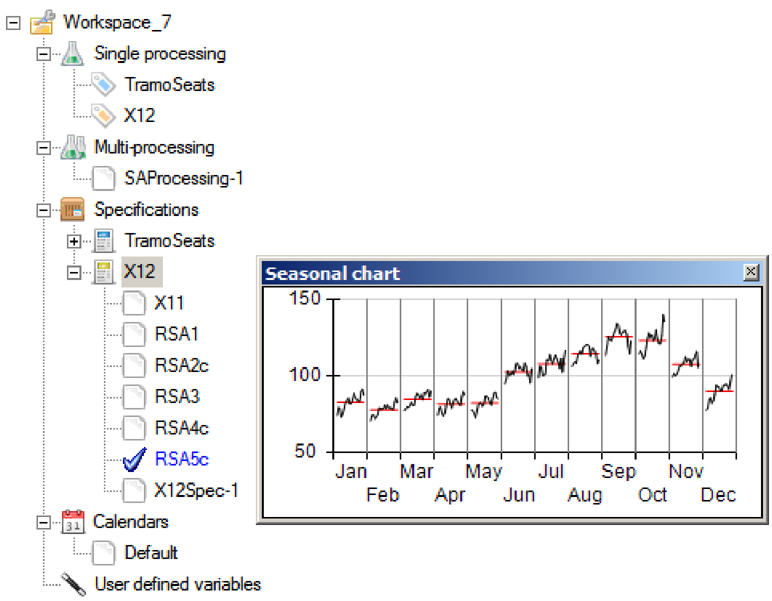



### 3.2.2 Tool window

*Tool* window offers the following options: **TS Properties**, **Chart**, **Growth Chart**, **Seasonal Chart**, **Spectral Analysis and Differencing**. First three of the above have been described in previous sections. Others are characterised below.

#### *3.2.2.1 Seasonal chart*

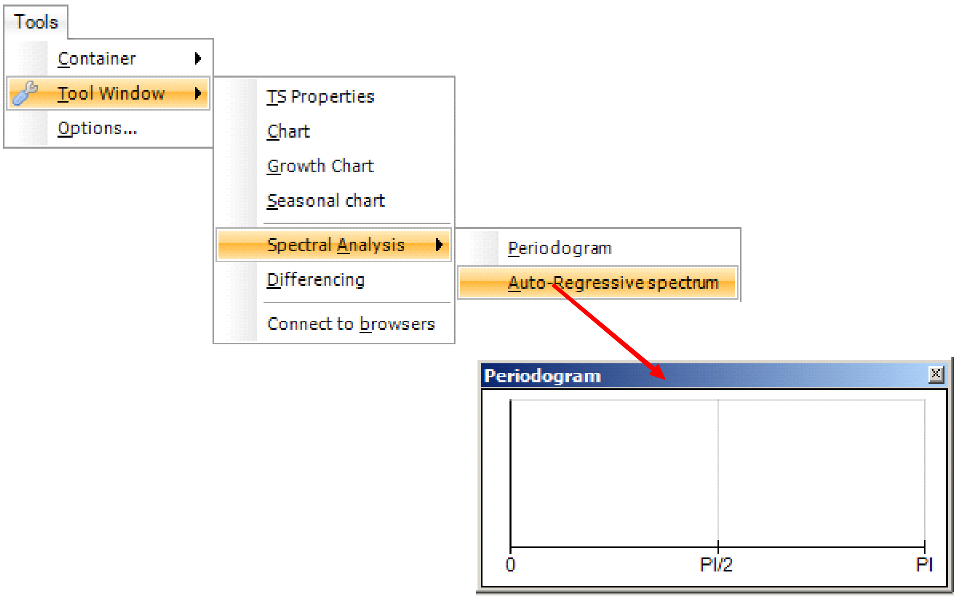
Seasonal chart presents the final estimation of the seasonal-irregular component and final seasonal factors for each of the period in time series (months or quarters). To calculate them Demetra+ uses the active specification (the one which is marked in the *Workspace menu*).



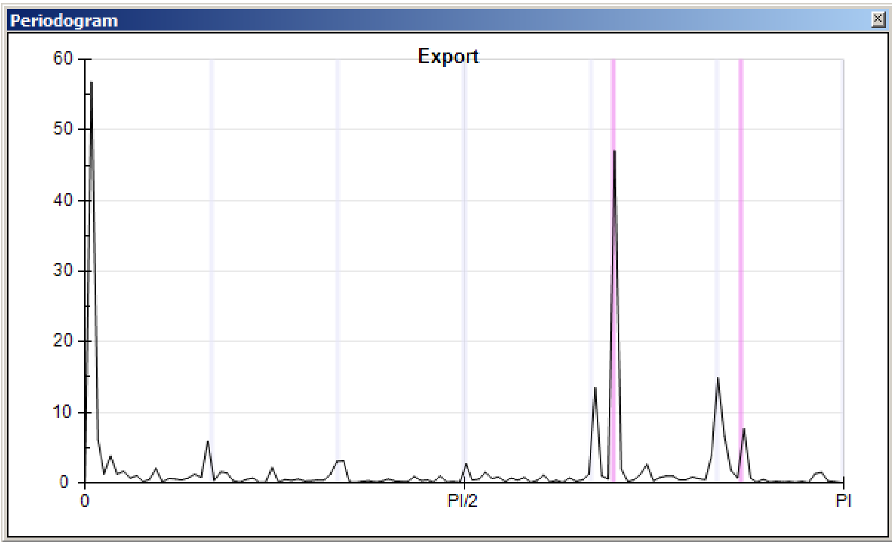

The curves visible on the chart represent the final seasonal factors and the straight line represents the average for these values in each period.

### 3.2.2.2 Spectral analysis

Demetra+ offers two spectral estimators – periodogram and autoregressive spectral estimator[7]. After choosing one of them from Tools menu the empty window is displayed.



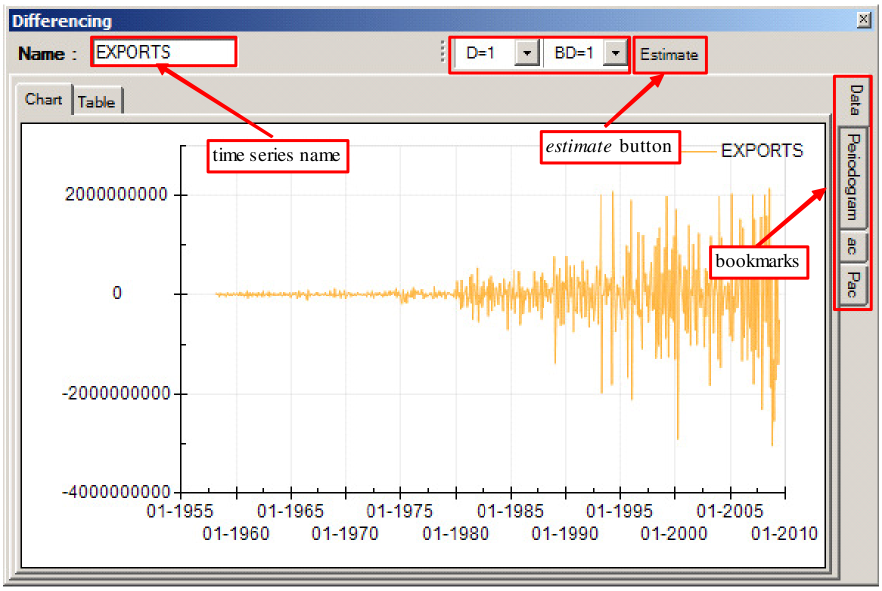

To calculate periodogram drag and drop a raw time series into the displayed window. The methodological note about spectral analysis is available further into this instruction.




### 3.2.2.3 Differencing

*Differencing* window gives the access not only to the table and spectral graphs but also to ACF and PACF functions for selected time series. To do it, the time series from the list should be dragged and dropped into *Name* box.

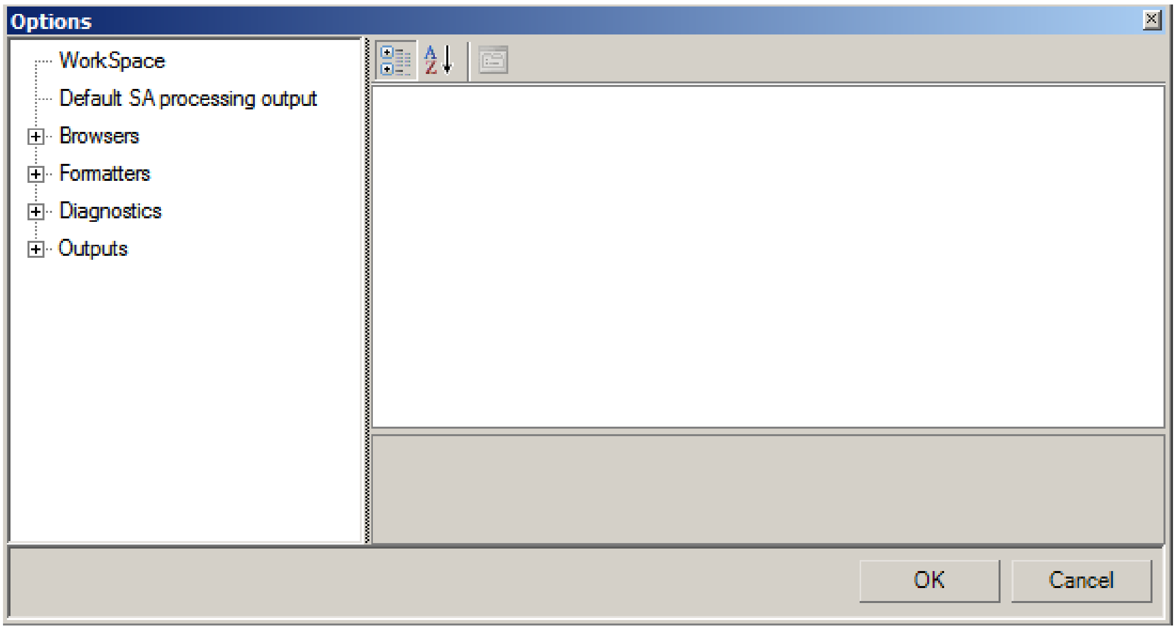
7 For more information see Annex.




Using the bookmarks on the right the user could switch other functions like periodogram and auto-regressive spectrum, autocorrelation function and partial autocorrelation.

### 3.2.2.4 Options

The window contains the default options used by the Demetra+.




The initial settings can be modified by the user. The menu includes:

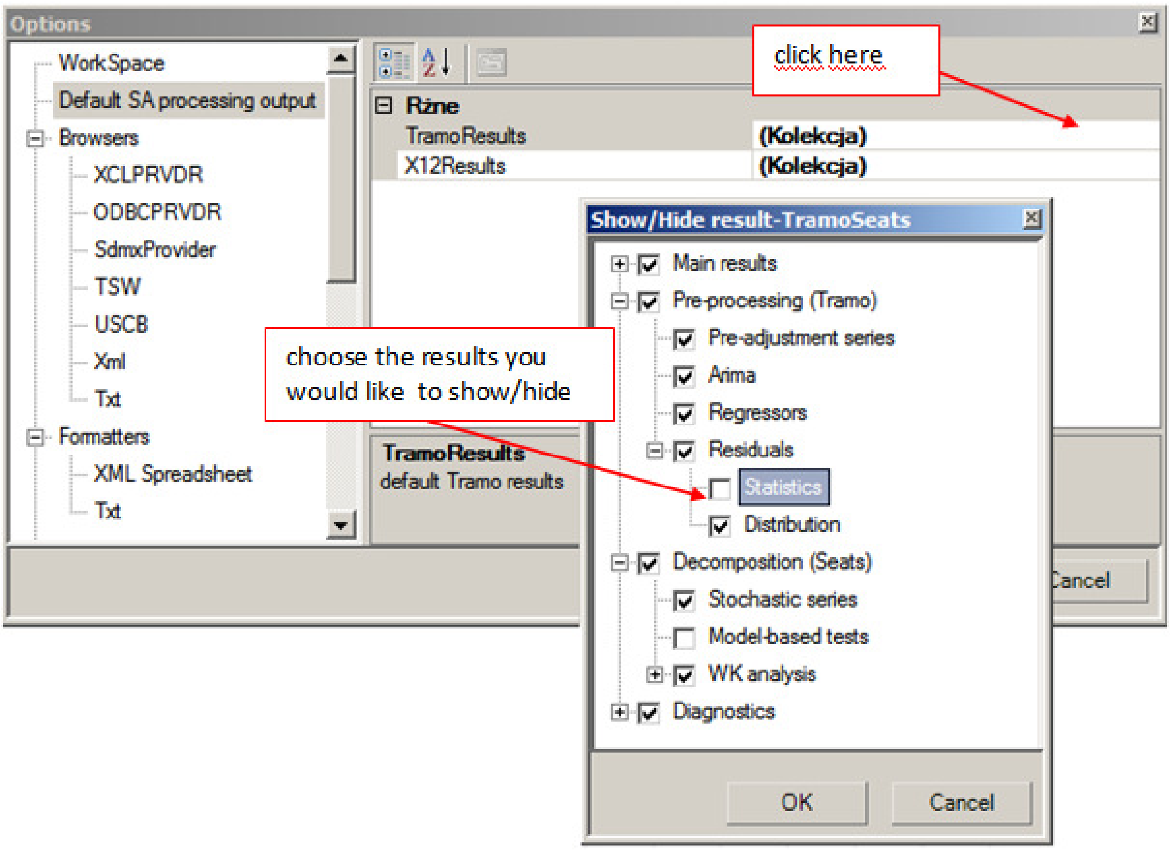
- setting for workspace,
- default processing output,
- settings for the browsers,

- formatters for txt and xml files,
- settings for presentation the diagnostic where the user can change the critical values and other parameters for diagnostic tests,
- outputs, where the folder that will contain the results is specified.

Some of those functions are discussed below.

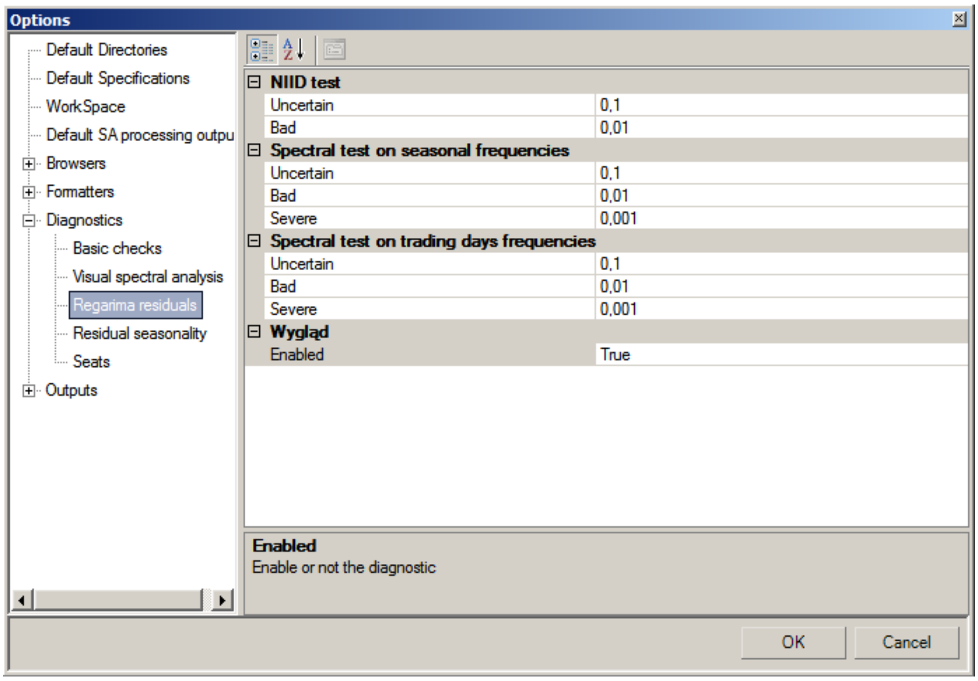
#### 3.2.2.4.1 Default SA processing output

The user can decide which parts of the results will be presented after SA processing. To do it, for each SA method the user can show or hide the items from the list of results. By default all the items are displayed after SA processing. The picture below presents that two diagnostic will not be visible in the SA results.




#### 3.2.2.4.2 Diagnostic

This part includes information about significance level for tests' result. For the spectral analysis the following settings are also included: threshold for identification of peaks, number of years (at the end of the series) considered in the spectral analysis, checking if the spectral peak appears on both SA series and irregular component. The default settings for the tests can be changed by the user.

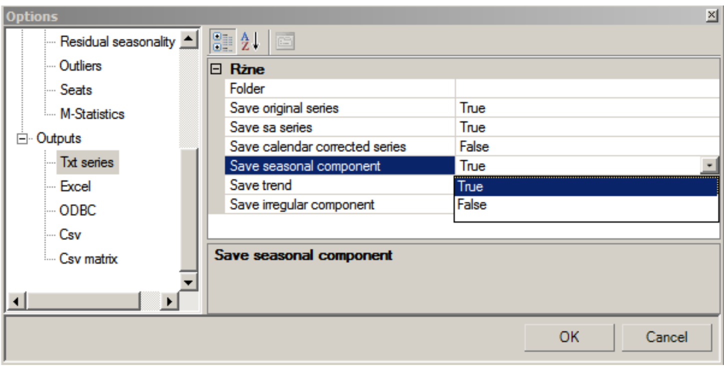



### 3.2.2.4.3 Outputs

This section enables to specify which output's items will be saved and folder in which Demetra+ saves the results. It is possible to save the results in the following formats: txt, excel, csv or send them to the database by ODBC.

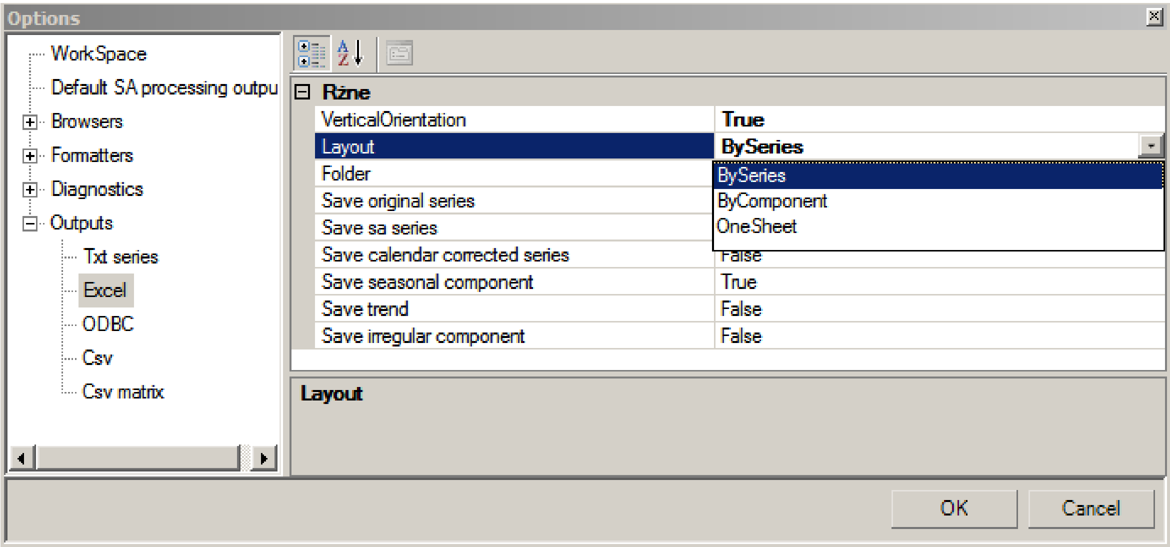
**TXT**

With the txt format the user can define the folder that will contain the results and the components that will be saved.

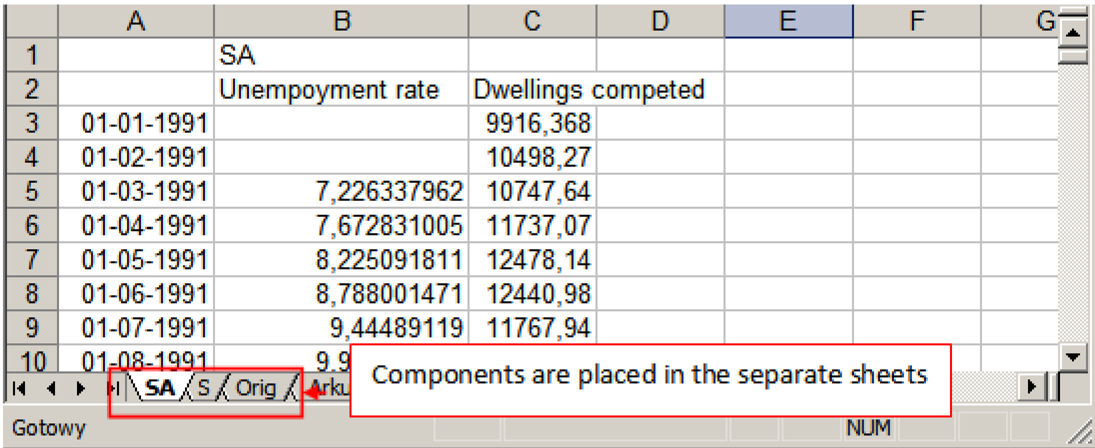



**XLS**

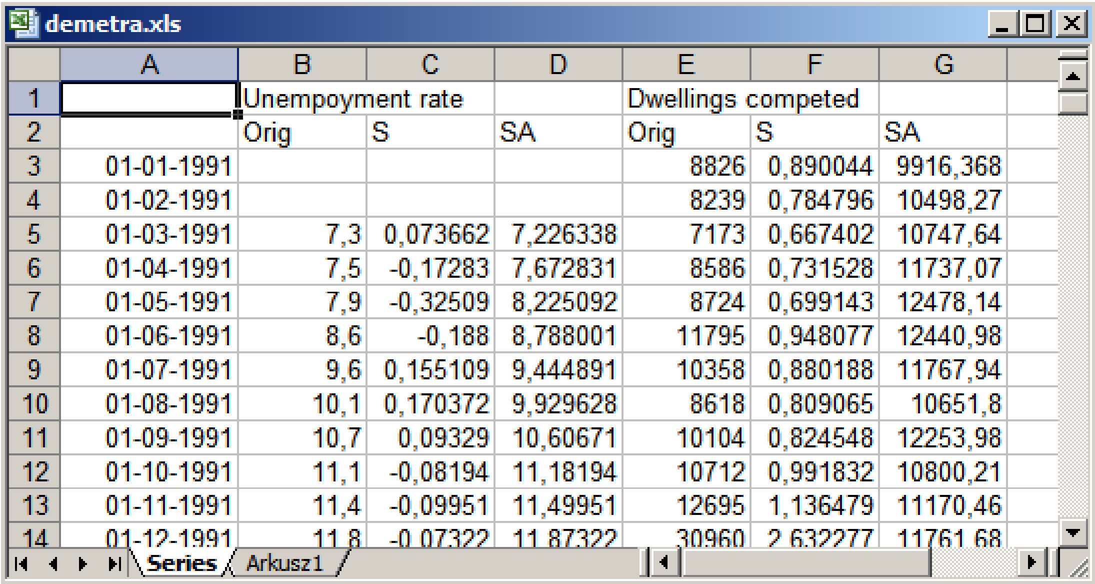
In addition to the options available for txt format, using xls format the user can specify the layout.




If the user will set the option layout to *ByCompenent* the output will be generated in the following way:




The option *OneSheet* will produce the following xls file:



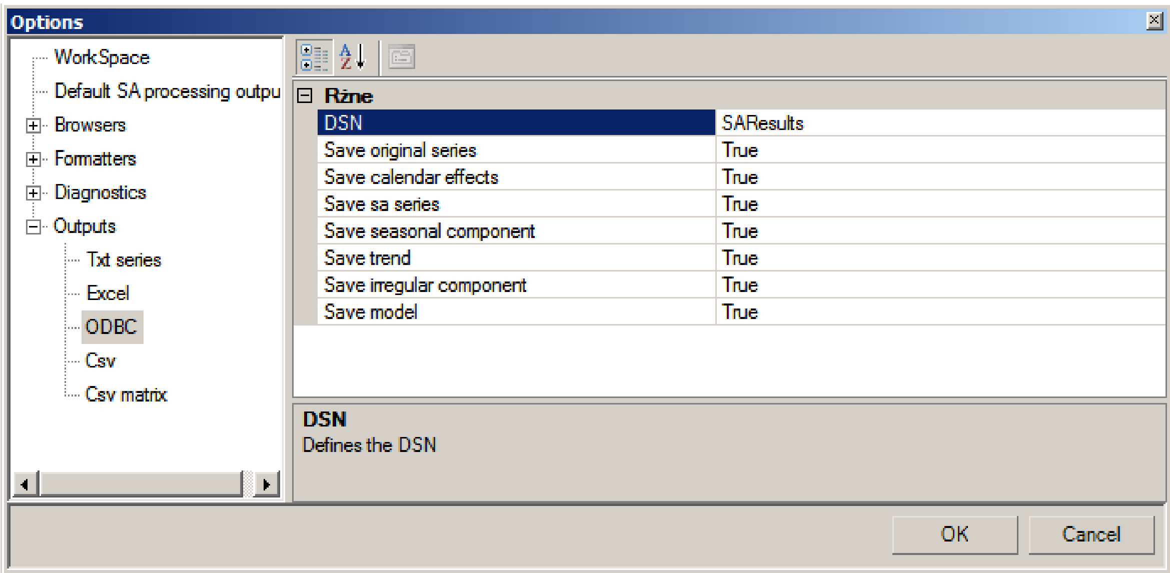
demetra.xls

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | | Unempoyment rate | | | Dwellings competed | | |
| 2 | | Orig | S | SA | Orig | S | SA |
| 3 | 01-01-1991 | | | | 8826 | 0,890044 | 9916,368 |
| 4 | 01-02-1991 | | | | 8239 | 0,784796 | 10498,27 |
| 5 | 01-03-1991 | 7,3 | 0,073662 | 7,226338 | 7173 | 0,667402 | 10747,64 |
| 6 | 01-04-1991 | 7,5 | -0,17283 | 7,672831 | 8586 | 0,731528 | 11737,07 |
| 7 | 01-05-1991 | 7,9 | -0,32509 | 8,225092 | 8724 | 0,699143 | 12478,14 |
| 8 | 01-06-1991 | 8,6 | -0,188 | 8,788001 | 11795 | 0,948077 | 12440,98 |
| 9 | 01-07-1991 | 9,6 | 0,155109 | 9,444891 | 10358 | 0,880188 | 11767,94 |
| 10 | 01-08-1991 | 10,1 | 0,170372 | 9,929628 | 8618 | 0,809065 | 10651,8 |
| 11 | 01-09-1991 | 10,7 | 0,09329 | 10,60671 | 10104 | 0,824548 | 12253,98 |
| 12 | 01-10-1991 | 11,1 | -0,08194 | 11,18194 | 10712 | 0,991832 | 10800,21 |
| 13 | 01-11-1991 | 11,4 | -0,09951 | 11,49951 | 12695 | 1,136479 | 11170,46 |
| 14 | 01-12-1991 | 11,8 | -0,07322 | 11,87322 | 30960 | 2,632277 | 11761,68 |

Series / Arkusz1

## ODBC

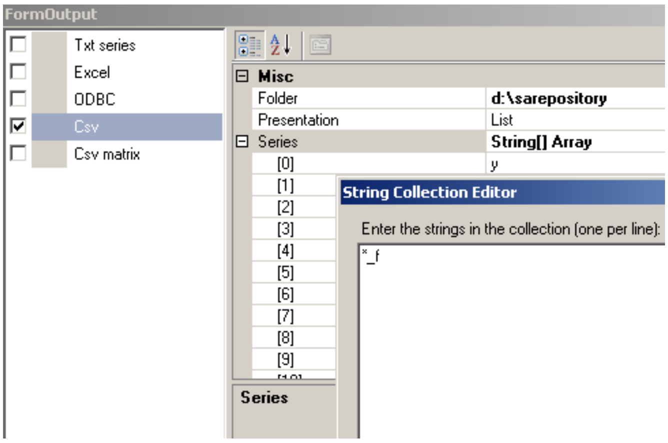
After choosing odbc option, the user should specify database source name (DSN). Needless to say, this database should be previously created. The user defines the components that will be sent to the database.




## CSV

Using the csv format, it is possible to generate for multi-processing documents a large number of time series generated by the models. Each file will contain, for all the series of the processing, a specific output (for instance, the calendar effects of all the series will be put together in one file). The different files will contain one item (row or column) for every series in the processing, even if that item is empty. The software can generate different layouts: the series can be presented in the form of horizontal or vertical tables (each row/column corresponding to the same period) or in the more compact form of horizontal lists of data.

The series must be introduced in the *String Collection Editor* (one code by row). The user can also use wildcards, in the usual way, to identify the series.




For example,

*_f

D*

y?

will generate all the forecasts, all the "D-tables" of X11 and the series "yc", "yl".

The different files will be stored as follows:

**<folder>\[<workspace name>]\<processing name>_<code>.csv**

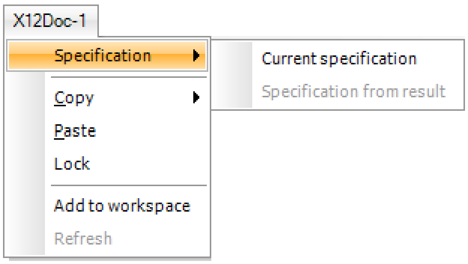
where:

- <**folder**> is specified by the user or the temporary folder if unspecified,
- <**workspace name**> is the workspace name (can be omitted),
- <**processing name**> is the name of the multi-processing,
- <**code**> has been defined above.

It should be noted that for multi-processing that don't belong to a workspace, the **[<workspace name>]\<processing name>** sequence is replaced by "**demetra**".

## 3.3 X-12 Doc

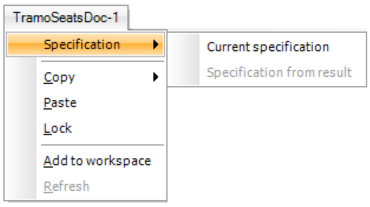
This item is visible in the application's menu when X-12 seasonal adjustment was previously done and after that it was activated by the user.




For a detailed description of the X12 specifications, you should refer to the "Demetra_Spec.docx" document.

## 3.4 TramoSeats Doc

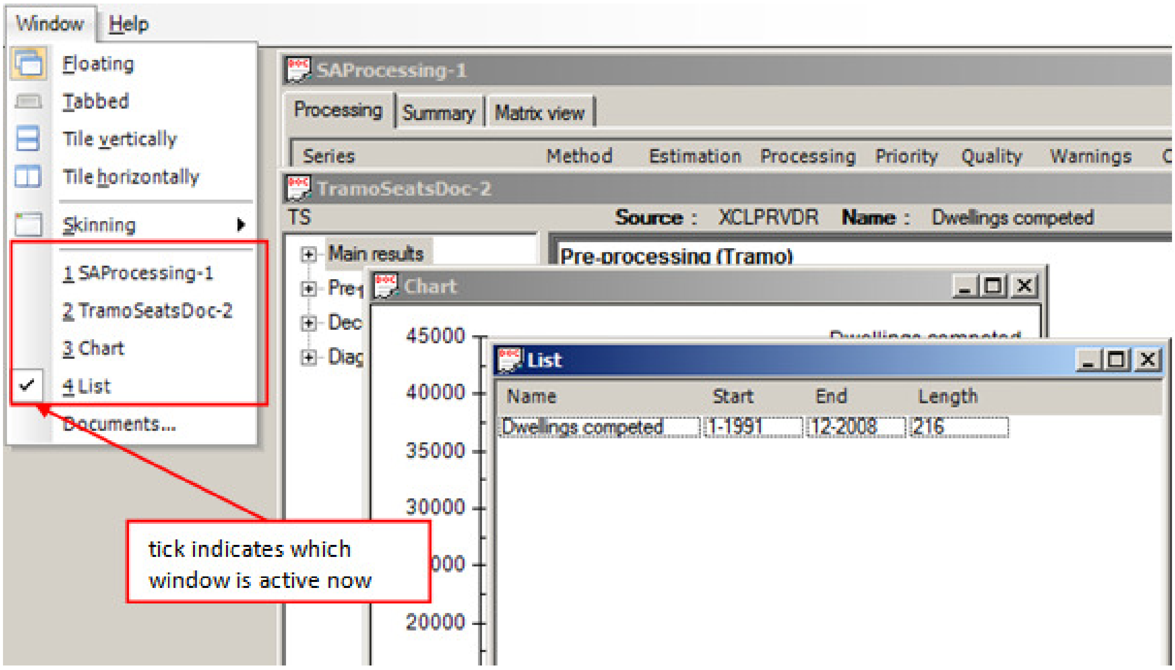
This item is added to the application's menu when seasonal adjustment using TramoSeats method was previously done and after that it was activated by the user. This item offers the similar options set as the X12Doc.



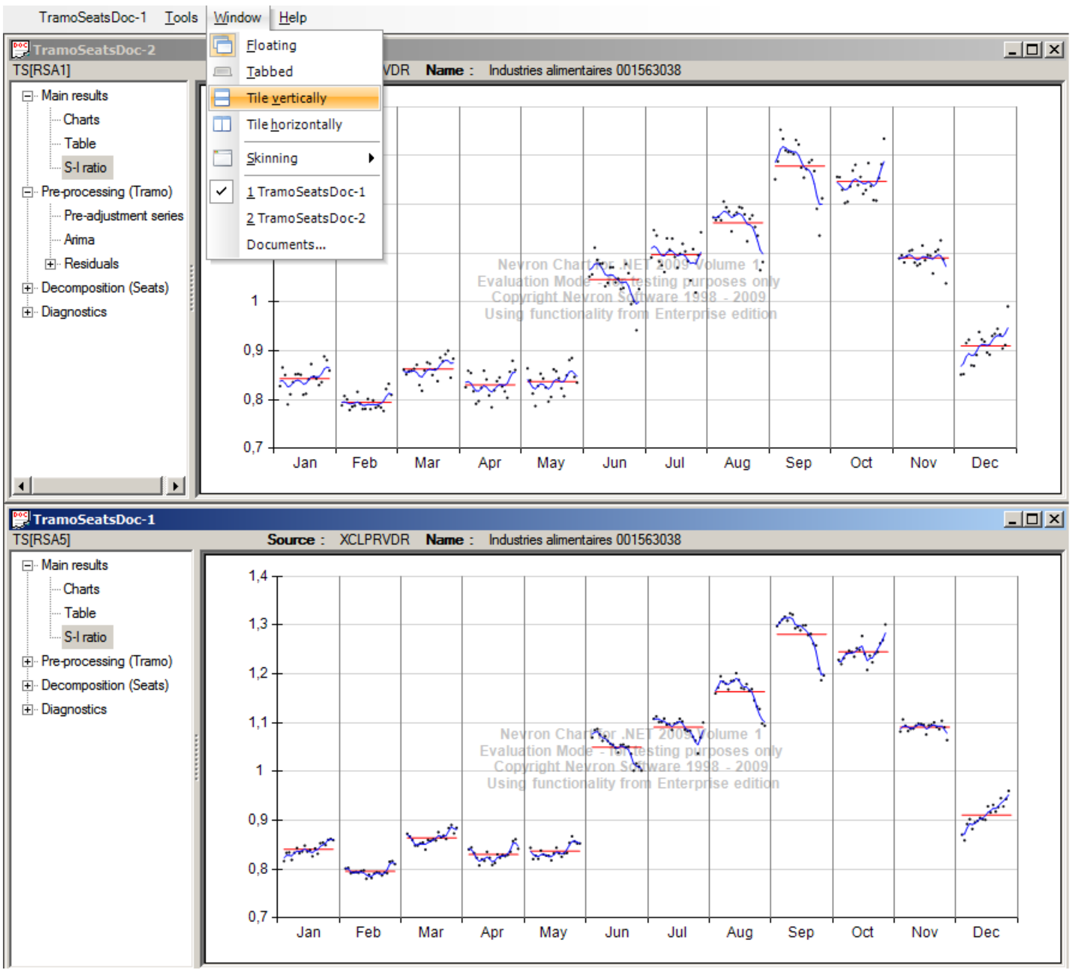

## 3.5 Window menu

Window menu offers the following functions:

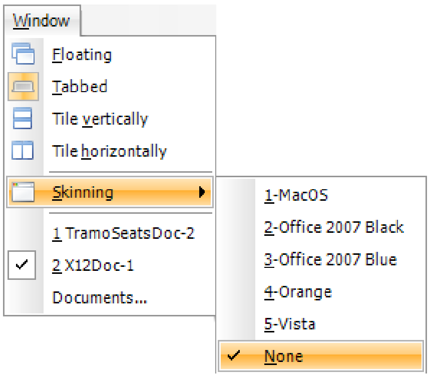
- **Floating** - show additional information while keeping the user in the same window,
- **Tabbed** - arranges all windows in central zone as tabs,
- **Tile vertically** -arranges all windows in central zone vertically,
- **Tile horizontally** -arranges all windows in central zone vertically,
- **Skinning** - allows to custom graphical appearance of Demetra+,
- **Documents,**
- List of windows' names in the central panel. This list is dynamically updated when the user opens/close some windows. On the example below four items are available. The one which is active is marked.




As an example, the following chart presents the comparison of the results for **Tile vertically** option.

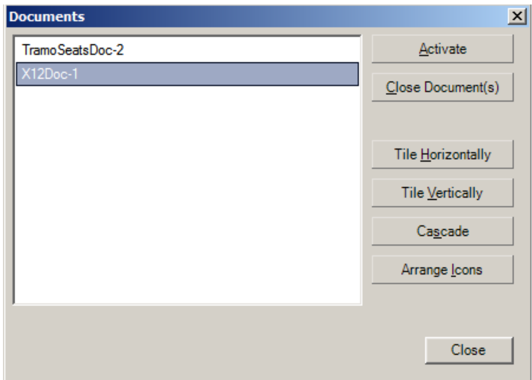


Demetra+ offers six different skinning:




The window menu includes also the seasonal adjustment processing done and not closed by the user during the current session.

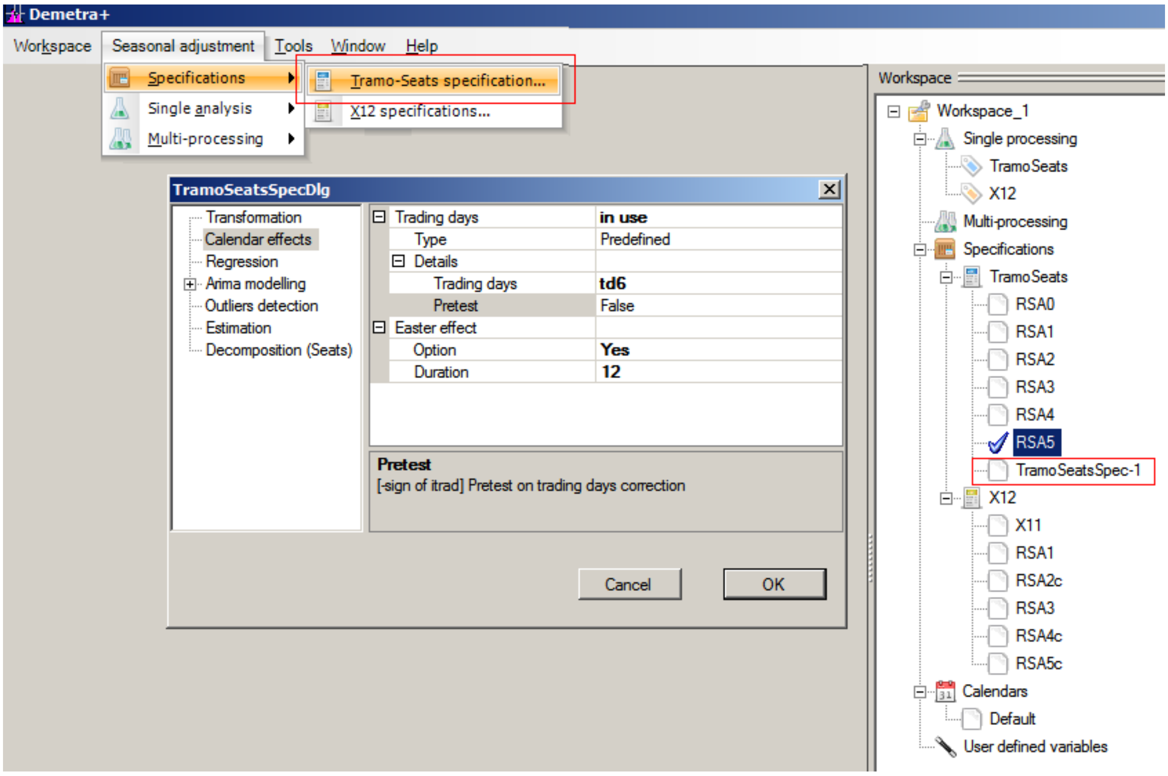
**Documents** option offers some additional options helpful for organising windows.



# 4 Seasonal adjustment

Demetra+ provides two methods of seasonal adjustment: TramoSeats and X12. For both methods a list of pre-defined specifications is available (using the naming conventions of TramoSeats). This list contains the most commonly used specification for seasonal adjustment. The description of the settings is available in the Annex. The default specifications appear in the Workspace tree. The users are strongly recommended to start their analysis - as explained below - with one of those specifications (usually RSA4(c) or RSA5(c)) and to change afterwards some of the options, if need be.

For more advanced users Demetra+ offers an opportunity to create the new specifications and add them to the list. This could be done by choosing the Seasonal Adjustment item of the main menu and clicking the Specifications sub-menu. In the next step the user should make a choice between **TramoSeats specification…** and **X12 specification…**. After the user has chosen all the suitable options in the Specifications dialog box, the new specification is automatically sent to the corresponding node of the Workspace. The new specification will be saved with the workspace, for future use. It can be used exactly like any predefined specification.




Next two sections contain valuable information about the specifications. The description of X12 specification is presented in 4.2 and the description of TS specification is presented in 4.3.

Demetra+ is able to perform seasonal adjustment for one single time series as well as for the whole set of time series. First option is called single processing (see 4.1) and is used for detailed analysis of the time series. Second option, called multi-processing (see 4.2), is a convenient tool for mass production of seasonally adjusted time series.

## 4.1 X12 specifications

The X12 specification is - to a very large extent - organized following the different individual specs of the original program (taking into account that peripheral specifications or specifications related to diagnostics are handled in a different way).

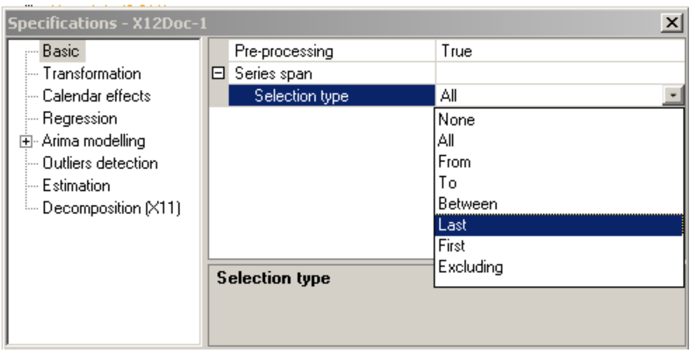
The different parts of the specification are presented in order in which they are displayed in the graphical interface of Demetra+. Details on the links between each item and its corresponding X12 spec/argument are provided in the following paragraphs. For an exact description of the different parameters, you should refer to the documentation of the original X12 program.

### 4.1.1 General description

| Item | X12 spec file | Meaning |
|---|---|---|
| Basic | *series* | General options for the processing |
| Transformation | *transform* | Transformation of the original series |
| Calendar effects | *regression* | Specification of the part of the regression related to calendar |
| Regression | *regression* | Specification of the part of the regression which is not specifically related to calendar |
| Automatic modeling | *automdl* | Automatic model identification |
| Arima | *arima* | Arima modeling |
| Outliers detection | *outlier* | Automatic outliers detection |
| Estimation | *estimate* | Options on the estimation procedure of the RegArima model |
| Decomposition (X11) | *x11 [forecast]* | X11 decomposition |

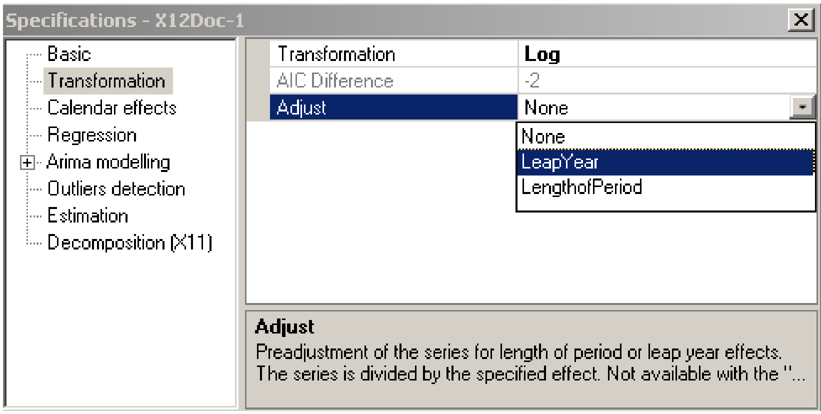
### 4.1.2 Basic

| Item | X12 | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| Pre-processing | | | Enable/Disable the other individuals specs, except X11 |
| Series span | *series* | *span* | Span (data interval) of the available time series used for the processing. The span can be computed dynamically on the series (for instance "Last 90 obs"). |



### 4.1.3 Transformation

| Item | X12 | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| Transformation | transform | function | Demetra+ accepts the following options:<br>• None,<br>• Log,<br>• Auto. |
| AIC Difference | transform | aicdiff | Disabled when the transformation is not set to "Auto" |
| Adjust | transform | adjust | Acceptable values:<br>• LeapYear - include a contrast variable for leap-year,<br>• LengthofPeriod - include length-of-month (or length-of-quarter) as a regression variable |




#### 4.1.3.1 Calendar effects

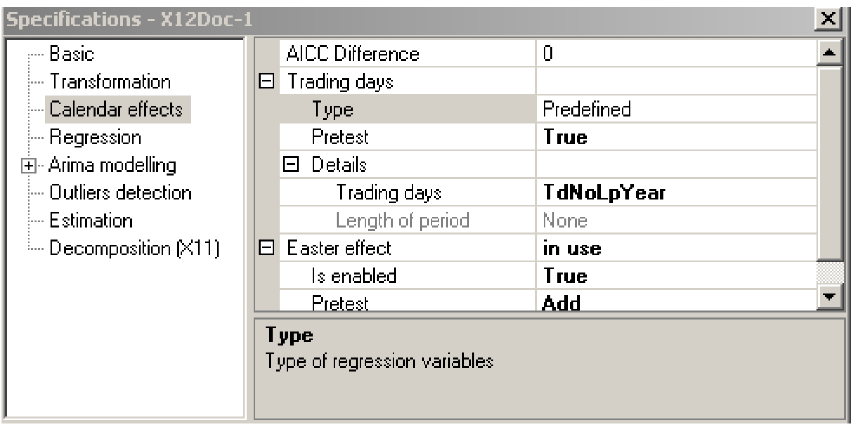
| Item | X12 | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| AICC Difference | *regression* | *aicdiff* | Demetra+ only considers pre-tests on regression variables related to calendar effects (trading days or moving holidays) |
| Type | - | - | The user can choose between:<br>• None,<br>• Predefined,<br>• Calendar,<br>• UserDefined.<br>"None" means that calendar effects will not be included in the regression.<br>"Predefined" means that default calendar will be used,<br>"Calendar" corresponds to the pre-defined trading days variables, modified to take into |

<table>
<tr><th rowspan="2">Item</th><th colspan="2">X12</th><th rowspan="2">Comments</th></tr>
<tr><th>Individual spec</th><th>Argument</th></tr>
<tr><td></td><td></td><td></td><td>account specific holidays. It means that the user should specify the type of trading days effect (td1,td2, td6 or td7) and chose calendar which will be used for holidays' estimation.<br>"UserDefined" is used when the user wants to specify in a free way his own trading day variables</td></tr>
<tr><td>Pretest</td><td>regression</td><td>aictest</td><td>Pretest the significance of the trading days regression variables using AICC statistics</td></tr>
<tr><td>Trading days (predefined type or calendar type)</td><td>regression</td><td>variables</td><td>Acceptable values:<br>• Td - include the six day-of-he week variables and a leap year effect,<br>• td1Coef - include the weekday-weekend contrast variable and a leap year effect,<br>• tdNoLpYear - include the six day-of-he week variables,<br>• td1NoLpYear - include the weekday-weekend contrast variable.<br>Some options can be disabled when the adjust option is used in the "transformation" specification</td></tr>
<tr><td>Length of period (predefined type or calendar type)</td><td>regression</td><td>variables</td><td>Acceptable values:<br>• LeapYear - include a contrast variable for leap-year,<br>• LengthofPeriod - include length-of-month (or length-of-quarter) as a regression variable .<br>Can be disabled when the adjust option is used in the "transformation" specification or with some trading days options.</td></tr>
<tr><td>Holidays (calendar type)</td><td>-</td><td>-</td><td>When the user chooses the "calendar" type for the trading days, he must specify the corresponding holidays. It should be noted that such a holiday must have been previously defined (see 3.1.1).</td></tr>
<tr><td>Items (UserDefined type)</td><td>regression</td><td>user, usertype= (...td...)</td><td>When the user chooses the "userdefined" type for the trading days, he must specify the corresponding variables. It should be noted that such variables must have been previously defined (see 3.1.2).</td></tr>
<tr><td>Easter (IsEnabled)</td><td>regression</td><td>variables and/or aictest</td><td>The option enables the user to estimate the Easter effect in tree different ways.<br>The user can choose between tree pre-test options:<br>• Add,<br>• Remove,</td></tr>
</table>

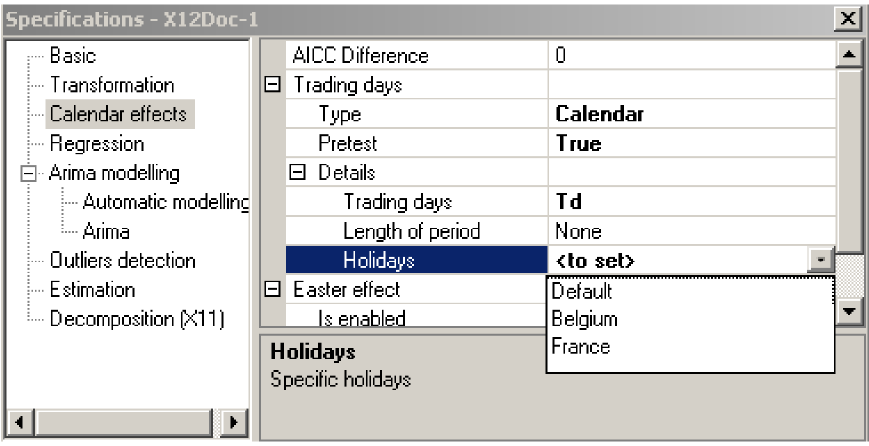
| Item | X12 | | Comments |
|---|---|---|---|
| | **Individual spec** | **Argument** | |
| | | | • None.<br>Trading days and holiday adjustments may be obtained from RegArima part or from irregular regression models.<br>When the user chooses the "Add", "easter" is only added in the "variables" part of the regression spec. An automatic identification of the Easter length (between 1, 8 and 15 days) is executed.<br>When he chooses the "Remove", "easter" is added in the "variables" an in the "aictest" parts of the regression spec. The specified length of the Easter effect is used.<br>When he chooses the "None", "easter" is only added in the "variables" of the regression spec. The length of the Easter effect specified by the user is used. The length of the Easter effect can range from 1 to 20 days.<br>It should be noted that the "Length" option is hidden when the "Add" pre-test option is active. |
| Pretest | *regression* | *aictest* | Pretest the significance of the Easter regression variables using AICC statistics |
| Length | *regression* | *easter[w]* | Duration of the Easter effect (w parameter of the "easter" variable. The parameter is active if the *aictest*=None |

The current version of Demetra+ doesn't allow the use of stock trading days. Regression variables for the handling of Labor Day and of Thanksgiving are not handled.

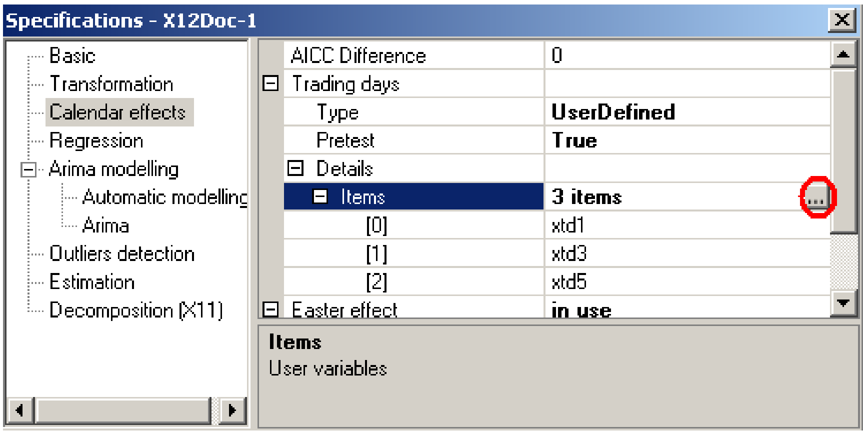
Example: predefined trading days



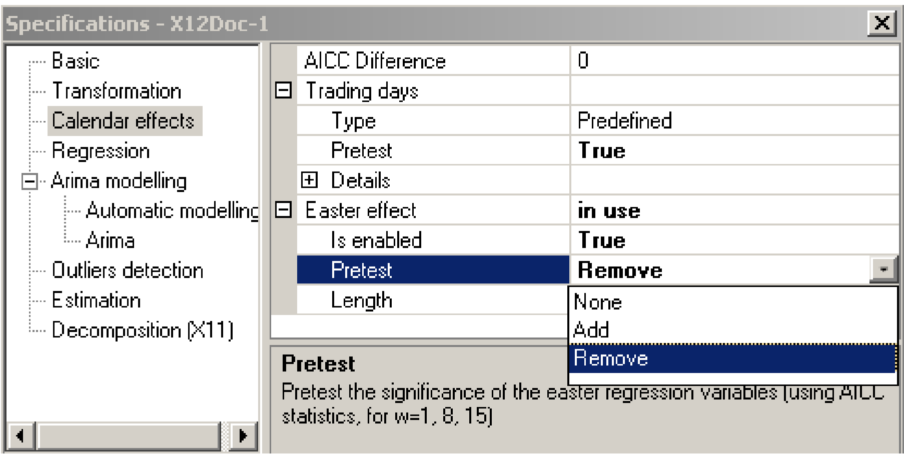
Example: calendar trading days




Example: user-defined trading days




Example: Easter effect

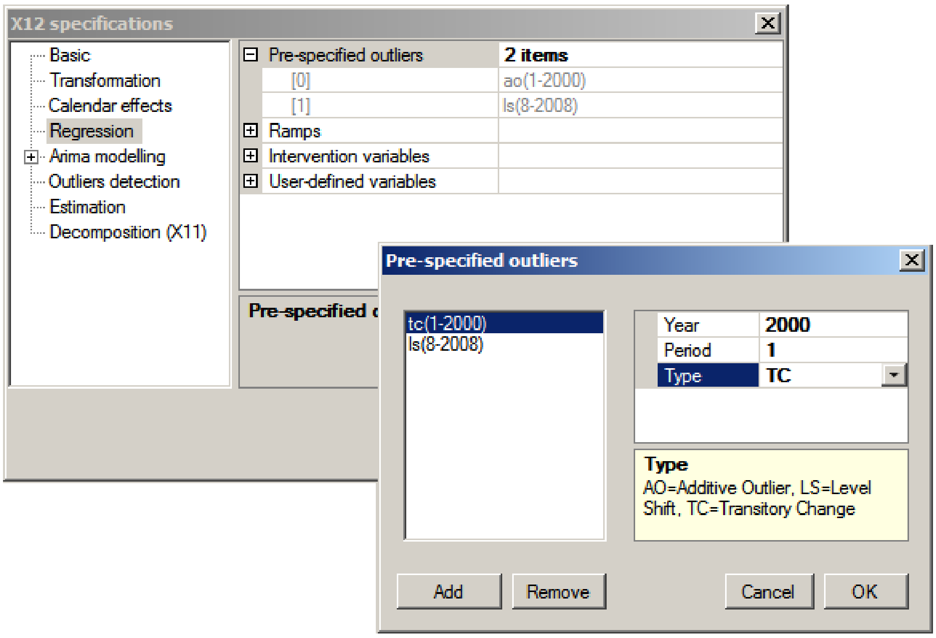


### 4.1.3.2 Regression

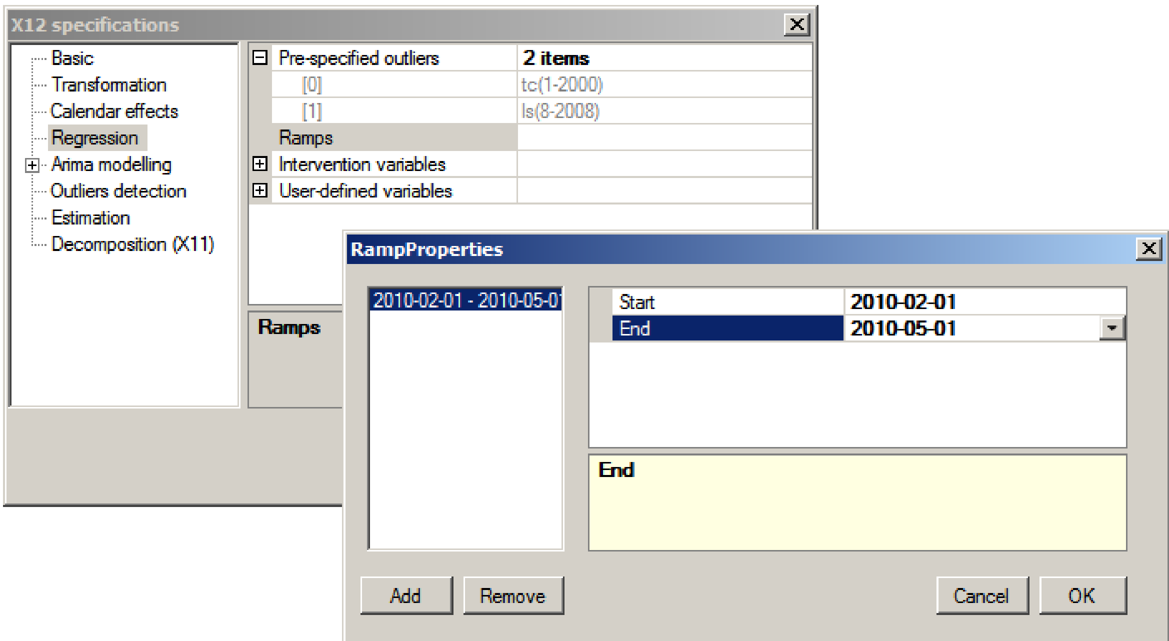
| Item | X12 | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| Pre-specified outliers | regression | variables | User-defined outliers used where prior knowledge suggest such effects at known time points:<br>• aoyyyy.pp – additive, point outlier which occurred in a given date (AO),<br>• lsyyyy.pp - regression variable for a constant level shift beginning on the given date (LS),<br>• tcyyyy.pp – regression variable for a temporary level change beginning on the given date (TC).<br>Seasonal outliers are not supported.<br>Pre-specified outliers are simple forms of intervention variables. |
| Ramps | regression | variables | Ramp effect which begins and ends on a given dates ( rpyyyy.pp-zzzz.qq). All dates of the ramps must occur within the time series. Ramps can overlap other rams, additive and level shifts outliers. |
| Intervention variables | regression | variables | No corresponding X12 arguments. The intervention variables are defined as in Tramo. Following their definition, these effects are special events known a-priori (strikes, devaluations, political evens, and so on). Intervention variables are modeled as any possible sequence of ones and zeros, on which some operators may be applied. This option enables the user to define four types of intervention variables:<br>• Dummy variables,<br>• Any possible sequence of ones and zeros,<br>• $\frac{1}{(1-\delta B)}$ of any sequence of ones and zeros ($0<\delta(Delta)\le 1$),<br>• $\frac{1}{(1-\delta_s B^s)}$ of any sequence of ones and zeros ($0<\delta_s(DeltaS)\le 1$)[8]. |
| User-defined variables | regression | user, usertype | The user-defined variables effect can be assigned to the trend, irregular of holiday or can exist as an additional component (option “None”). For those variables the user can specify lags. |

8 See BOX, G. E. P. and TIAO, G. C. (1975).

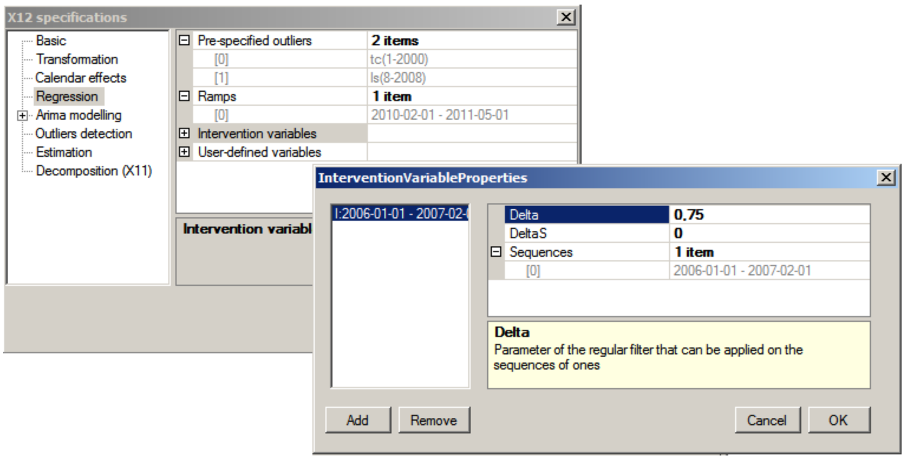
Example: Pre-specified outliers



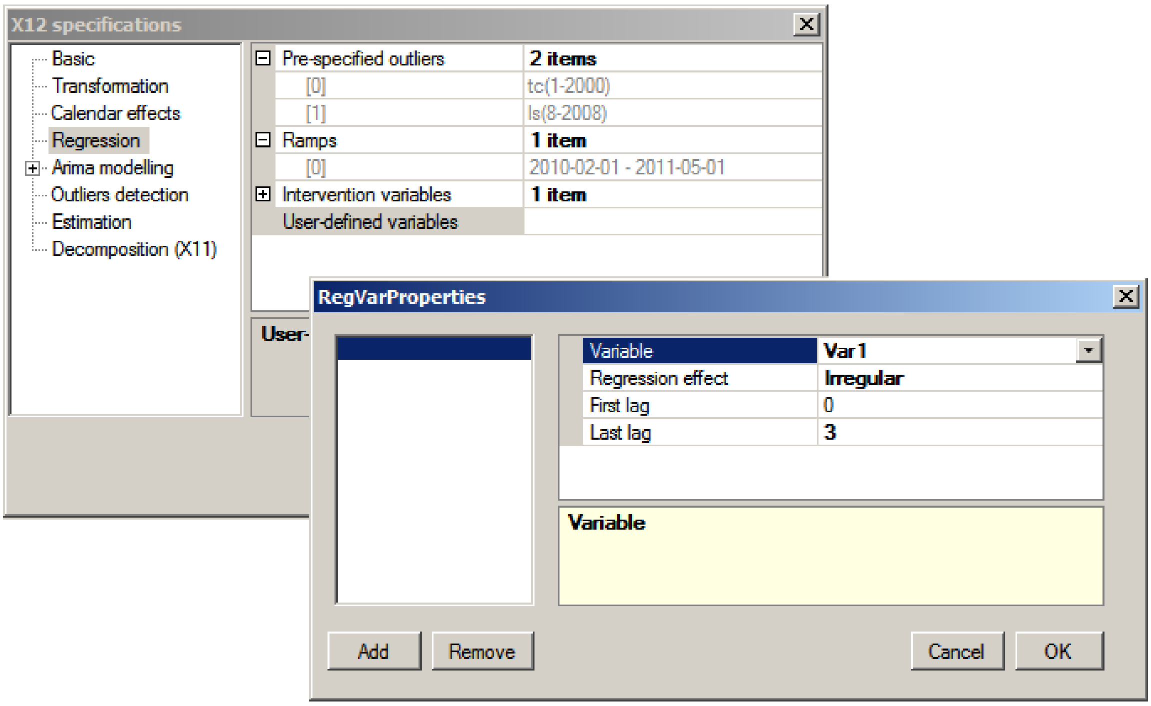

Example: Ramps



Example: Intervention variables



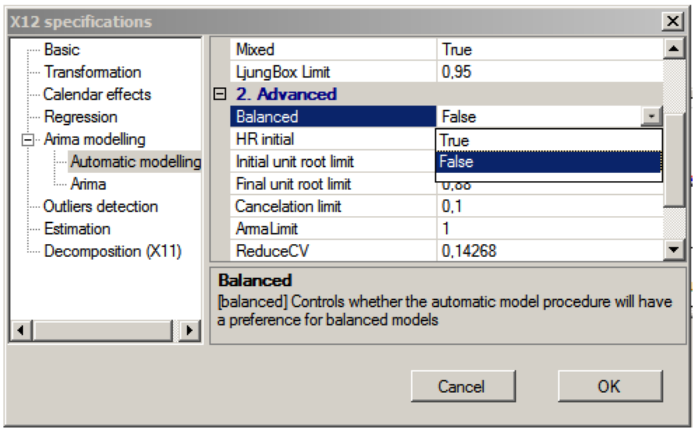

Example: User-defined variables



#### *4.1.3.3 Automatic modelling*

| Item | X12 | | Comments |
|---|---|---|---|
| | **Individual spec** | **Argument** | |
| IsEnabled | *automdl* | | Presence or not of the automdl individual spec |
| Accept default | *automdl* | *acceptdefault* | Controls weather the default model is chosen if the Ljung-Box Q statistics for its model residuals is acceptable |
| Check Mu | *automdl* | *checkmu* | Controls weather the automatic model selection procedure will check for the significance of a constant term |
| Mixed | *automdl* | *mixed* | Controls weather ARIMA models with nonseasonal AR and MA terms will be considered in the automodel |
| LjungBox limit | *automdl* | *ljungboxlimit* | Acceptance criterion for confident of the Ljung-Box Q statistic |
| Balanced | *automdl* | *balanced* | Controls weather the automatic model procedure will have a preference for balanced model |
| HR initial | *automdl* | *hrinitial* | Control weather Hannan-Rissanen[9] estimation is done before exact maximum likelihood estimation to provide initial values |
| Initial unit root limit | *automdl* | | Threshold value for the initial unit root test in the automatic differencing procedure |
| Final unit root limit | *automdl* | | Threshold value for the final unit root test in the automatic differencing procedure. This value should be greater than one. |
| Cancelation limit | *automdl* | | Cancelation limit for AR and MA roots |
| ArmaLimit | *automdl* | *armalimit* | Threshold value for t-statistics of ARIMA coefficients used for final test of model parsimony |
| ReduceCV | *automdl* | *reducecv* | The percentage by which the outlier critical value will be reduced when an identical model is found to have a Ljung-Box statistic with an unacceptable confidence coefficient |
| Reduce SE | *automdl* | unavailable | Percentage reduction of SE |
| Unit root limit | *automdl* | *urfinal* | Unit root limit for final model. Should be >1 |

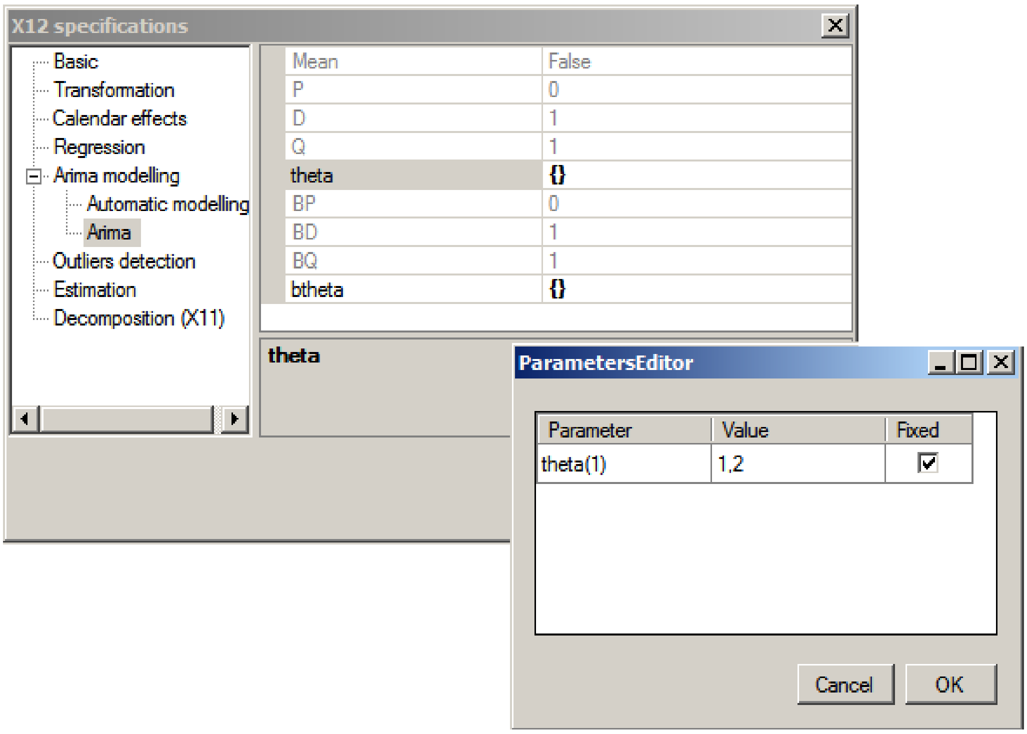
[9] According to GÓMEZ, V., and MARAVALL, A. (2001),the Hannan-Rissanen method is a penalty function method based on BIC (Bayesian Information Criterion) where the estimates of ARMA model parameters are computed by means of linear regressions. These estimates are computationally cheap and have similar properties to those obtained by Maximum Likelihood.




### 4.1.3.4 Arima

Options included in this section are active only if *IsEnabled* parameter from Automatic modeling section is set to false.

| Item | X12 | | Comments |
|---|---|---|---|
| | **Individual spec** | **Argument** | |
| Mean | regression | variables | It is considered that the mean is part of the Arima model (it highly depends on the chosen model). |
| P, D, Q, BP, BD, BQ | arima | model | Only "Box-Jenkins" SARIMA models (p d q)(bp bd bq) is considered |
| theta, btheta, phi, bphi | arima | | The coefficients are defined using the convention used in TramoSeats. It means that they are the opposite of the coefficients used in X12 |

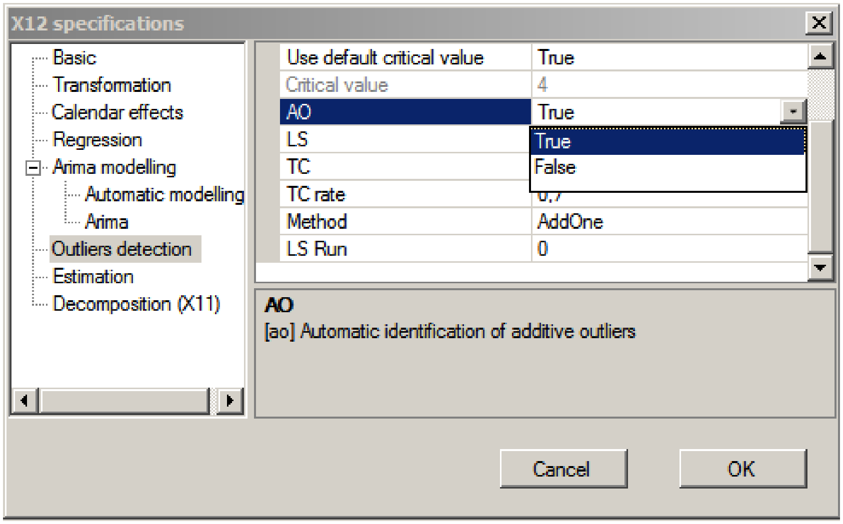


### 4.1.3.5 Outliers detection

Both X-12ARIMA and TramoSeats detect outliers, which are defined as the abrupt changes that cannot be explained by the underlying normality of the ARIMA model. Three outliers' types are detected: additive outlier (AO) which affects an isolated observation; level shifts (LS), which implies a step change in the mean level of the series; transitory change (TC), which describes a temporary effect on the level of series after a certain point in time[10].

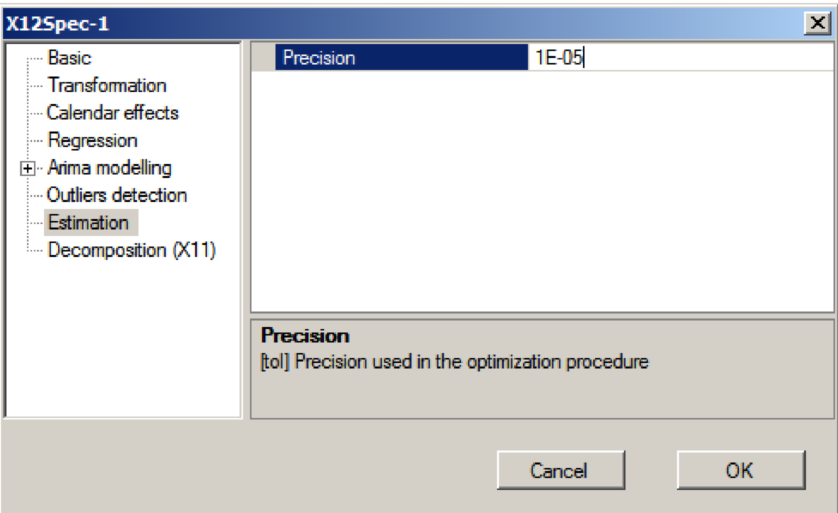
| Item | X12 | | Comments |
|---|---|---|---|
| | **Individual spec** | **Argument** | |
| IsEnabled | *outlier* | | Presence or not of the outlier individual spec |
| Outliers detection span | *outlier* | *span* | Span used for the outlier detection. The span can be computed dynamically on the series (for instance "Excluding last 12 obs") |
| Use default critical value | *outlier* | *critical* | When "Use default critical value" is false, the procedure uses the critical value mentioned in the specification. Otherwise, the default is used (the first case corresponds to "critical = xxx"; the second corresponds to a spec without the critical argument). It should be noted that it is not possible to define different critical values for different outliers' types |
| Critical value | *outlier* | *critical* | Critical value used in the outliers detection procedure |
| AO | *outlier* | *ao* | Automatic identification of additive outliers |
| LS | *outlier* | *ls* | Automatic identification of level shifts |
| TC | *outlier* | *tc* | Automatic identification of transitory changes |
| TC rate | *outlier* | *tcrate* | Rate of decay for transitory change outlier regressor |
| Method | *outlier* | *method* | Determines how the program successively adds detected outliers to the model (could be add one by one (the outliers with the highest/insignificant t-statistic is added/removed at one time and the ARIMA model estimated and so on) or add all outliers together (all the significant/insignificant outliers are added/removed at once and the ARIMA model estimated and so on) |
| LS Run | *outlier* | *lsrun* | Compute t-statistic to test null hypotheses that each run of $n$ *lsrun* successive level shifts cancels to form a temporary level shift |

10 KAISER, R., and MARAVALL, A. (2000).



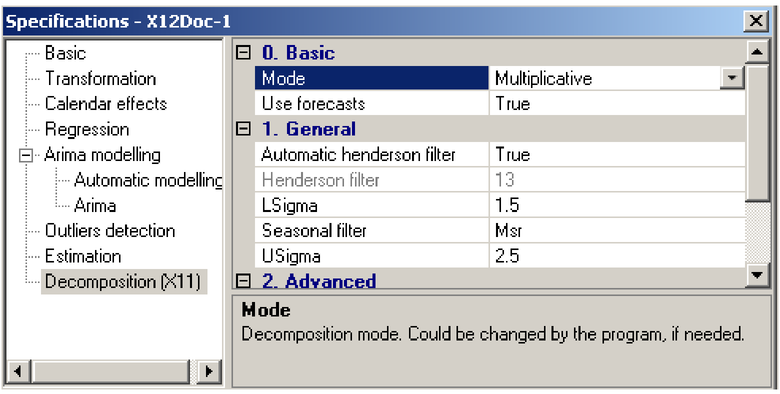

#### 4.1.3.6 Estimation

| Item | X12 | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| Precision | *estimate* | *tol* | Precision used in the optimization procedure |



### 4.1.3.7 Decomposition (X11)

| Item | X12 | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| Mode | *x11* | *mode* | Only additive mode. Pseudo-additive mode is not supported |
| Use forecasts | *forecast* | *maxlead* | When UseForecasts is false, maxlead is set to 0 and the x11 procedure doesn't use any model-based forecasts. Otherwise, the forecasts of the RegArima model (default airline model if the user doesn't use pre-processing - see basic options - ) is used to extend the series |
| LSigma | *x11* | *sigmalim* | First parameter of *sigmalim* - lower sigma boundary for the detection of the extreme values |
| USigma | *x11* | *sigmalim* | Second parameter of *sigmalim* - uppersigma boundary for the detection of the extreme values |
| Seasonal filter | *x11* | *seasonalma* | Specifies which seasonal moving average (seasonal filter) will be used to estimate the seasonal factors |
| Details on seasonal filters | *x11* | unavailable | Details on specific *seasonalma* for the different periods |
| Automatic Henderson filter | *x11* | *trendma* | Automatic selection of the Henderson filter is used when the corresponding item is selected. Otherwise, the length given by the user is used. |
| True 7 Term | *x11* | *true7term* | Specifies the end weights used for the seven term Henderson filter |
| Calendar Sigma | *x11* | *calendarsigma* | Specifies if standard errors used for extreme values detection and adjustment are computed separately for each calendar period (month, quarter) or separately for two complementary sets of calendar periods. |
| Sigma Vector | *x11* | unavailable | Specifies one of the two groups of periods for whose irregulars a group standard error will be calculated under the *calendarsigma*="Select" option |



## 4.2 TramoSeats specifications

The TramoSeats specification is based on the original program (taking into account that peripheral specifications or specifications related to diagnostics are handled in a different way).

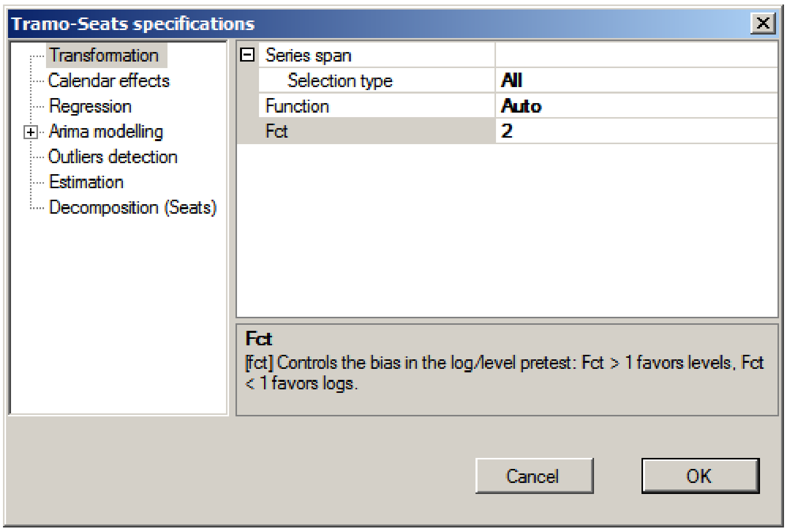
The different parts of the specification are presented in order in which they are displayed in the graphical interface of Demetra+. Details on the links between each item and its corresponding X12 spec/argument are provided in the following paragraphs. For an exact description of the different parameters, the user should refer to the documentation of the original TramoSeats program.

### 4.2.1 General description

| **Item** | **TramoSeats spec file** | **Meaning** |
|---|---|---|
| Transformation | *transform* | Transformation of the original series |
| Calendar effects | *regression* | Specification of the part of the regression related to calendar |
| Regression | *regression* | Specification of the part of the regression which is not specifically related to calendar |
| Automatic modelling | *automdl* | Automatic model identification |
| Arima | *arima* | Arima modelling |
| Outliers detection | *outlier* | Automatic outliers detection |
| Estimation | *estimate* | Options on the estimation procedure of the RegArima model |
| Decomposition (Seats) | *seats* | Seats decomposition |

### 4.2.2 Transformation

| **Item** | **TramoSeats** | | **Comments** |
|---|---|---|---|
| | **Individual spec** | **Argument** | |
| Series span | transform | *span* | Span used for the processing. The span can be computed dynamically on the series (for instance "Last 90 obs") |
| Function | transform | *lam* | Transformation of data: logarithm or none |
| Fct | transform | *fct* | Control the bias in the log/level pretest (the function is active if *lam=Auto)*. Fct > 1 favors levels, Fct < 1 favors logs |



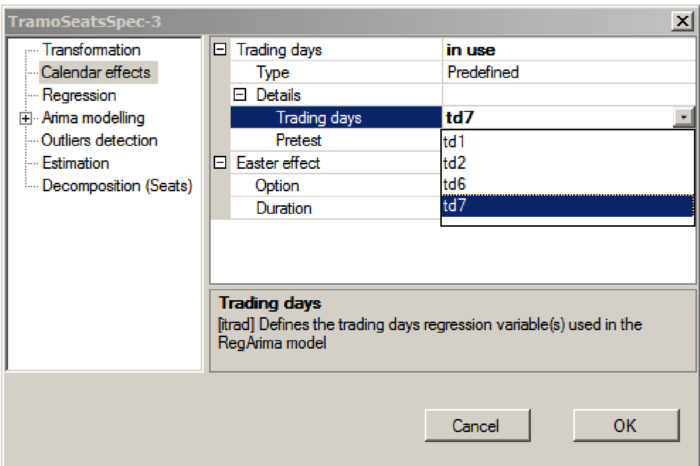

#### 4.2.2.1 Calendar effects

| Item | TramoSeats | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| Type | - | - | The user can choose between:<br>• None,<br>• Predefined,<br>• Calendar,<br>• UserDefined.<br>"None" means that calendar effects will not be included in the regression.<br>"Predefined" means that default calendar will be used,<br>"Calendar" corresponds to the pre-defined trading days variables, modified to take into account specific holidays. It means that the user should specify the type of trading days effect (td1,td2, td6 or td7) and chose calendar which will be used for holidays' estimation.<br>"UserDefined" is used when the user wants to specify in a free way his own trading day variables |
| Trading days (predefined type or calendar type) | *regression* | *variables* | Acceptable values:<br>• td1 - include the weekday-weekend contrast variable,<br>• td2 - include the weekday-weekend contrast variable and a leap year effect,<br>• td6 - include the six day-of-he week variables,<br>• td7 - include the six day-of-he week variables and a leap year effect,. |

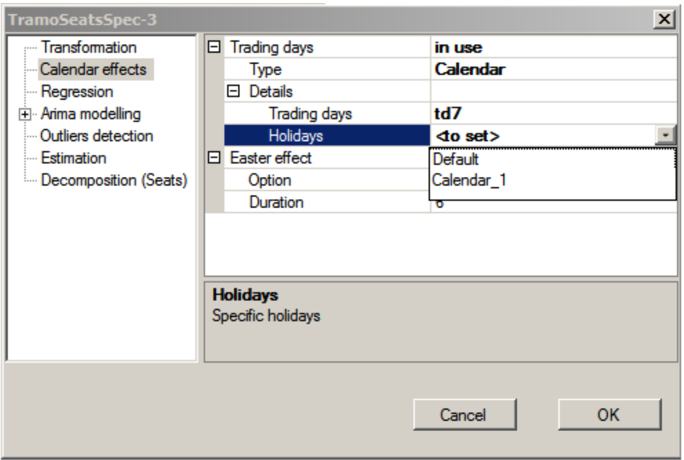
| Item | TramoSeats | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| Pretest | *regression* | *aictest* | Pretest the trading days correction. Option available for type=Predefined |
| Holidays (calendar type) | - | - | When the user chooses the "calendar" type for the trading days, he must specify the corresponding holidays. It should be noted that such a holiday must have been previously defined (see 3.1.1). |
| Items (UserDefined type) | *regression* | | When the user chooses for the trading days the type="UserDefined", he must specify the corresponding variables. It should be noted that such variables should have been previously defined (see 3.1.2). |
| Easter (IsEnabled) | *regression* | *variables and/or aictest* | The option enables the user to estimate the Easter effect in tree different ways.<br>The user can choose between:<br>• No<br>• Pretest<br>• Yes<br>"No" - a correction for Easter effect is not performed.<br>"Pretest" meant that Demetra+ tests for the necessity of a correction for Easter effect.<br>"Yes" – the correction for Easter effects is performed. For last two option the user can modify the default length of the Easter effect (default length is 6 days) |
| Length | *regression* | *easter[w]* | Duration of the Easter effect ( $w$ parameter of the "easter" variable |

The current version of Demetra+ doesn't allow the use of stock trading days. Regression variables for the handling of Labor Day and of Thanksgiving are not handled.

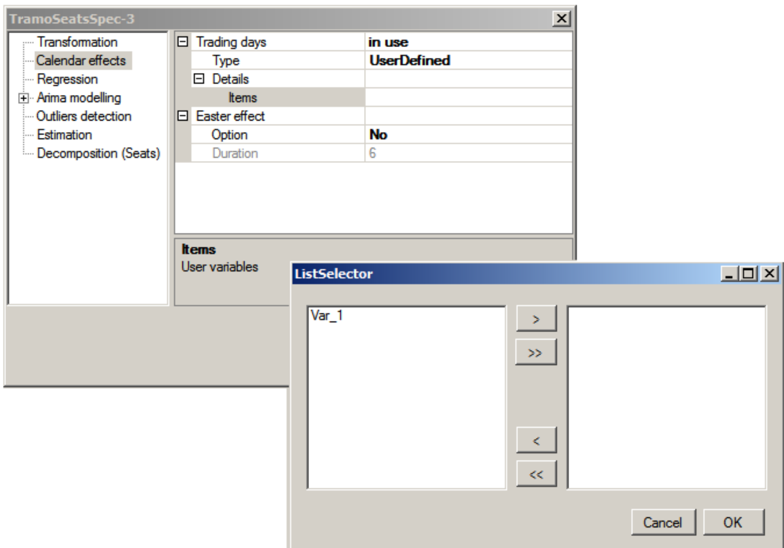
Example: predefined trading days



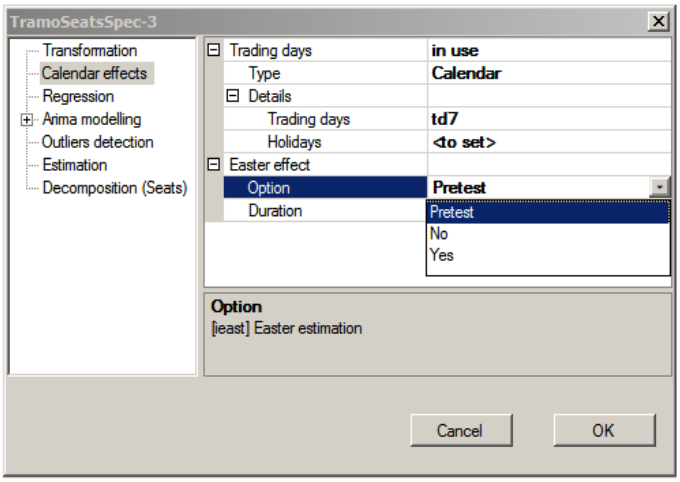
Example: calendar trading days




Example: user-defined trading days



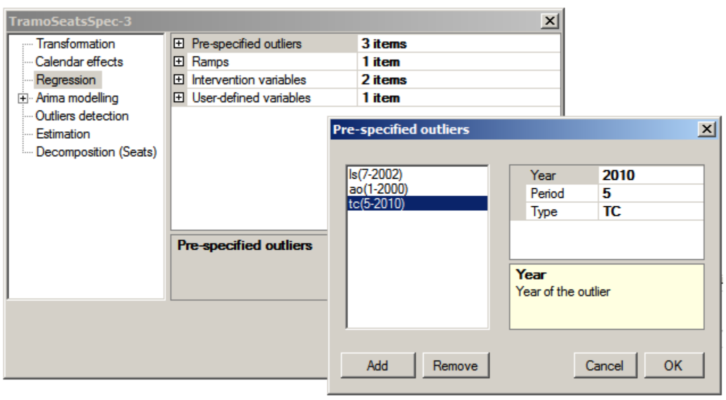

Example: Easter effect



### 4.2.2.2 Regression

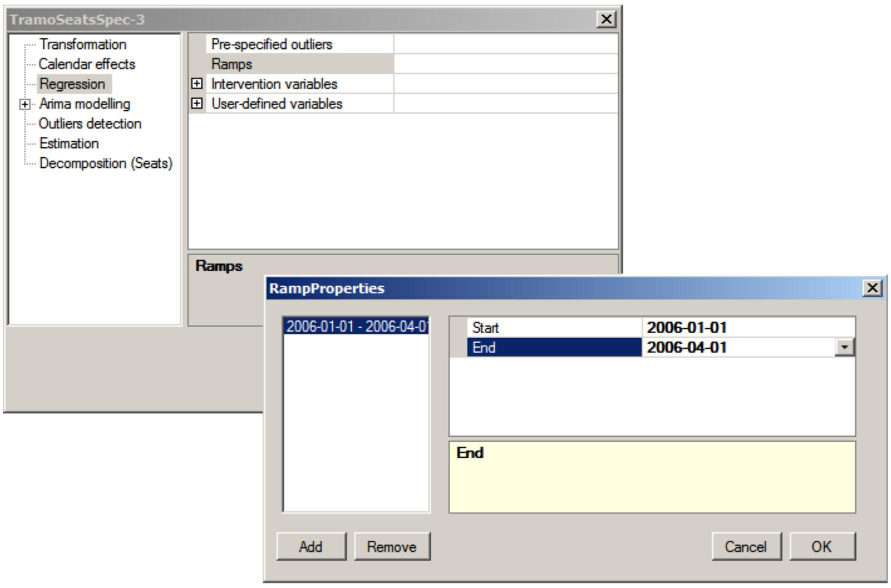
| Item | TramoSeats | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| Pre-specified outliers | regression | variables | Corresponds to the pre-specified outliers of TramoSeats:<br>• aoyyyy.pp,<br>• lsyyyy.pp,<br>• tcyyyy.pp. |
| Ramps | regression | variables | Corresponds to the ramps of TramoSeats: rpyyyy.pp-zzzz.qq |
| Intervention variables | regression | variables | Enables the user to define four types of intervention variables:<br>• Dummy variables,<br>• Any possible sequence of ones and zeros,<br>• $\frac{1}{(1-\delta B)}$ of any sequence of ones and zeros ($0 < \delta(Delta) \leq 1$),<br>• $\frac{1}{(1-\delta_s B^s)}$ of any sequence of ones and zeros ($0 < \delta_s(DeltaS) \leq 1$). |
| User-defined variables | regression | user, usertype | The user-defined variables are input by the user and can be considered as belonging to the trend, to the irregular component or to calendar effects (using the corresponding *ls, tc* and holiday user types). For practical considerations, seasonal effects are currently not supported. The user can specify the structure of the lags[11]. |

Example: Pre-specified outliers

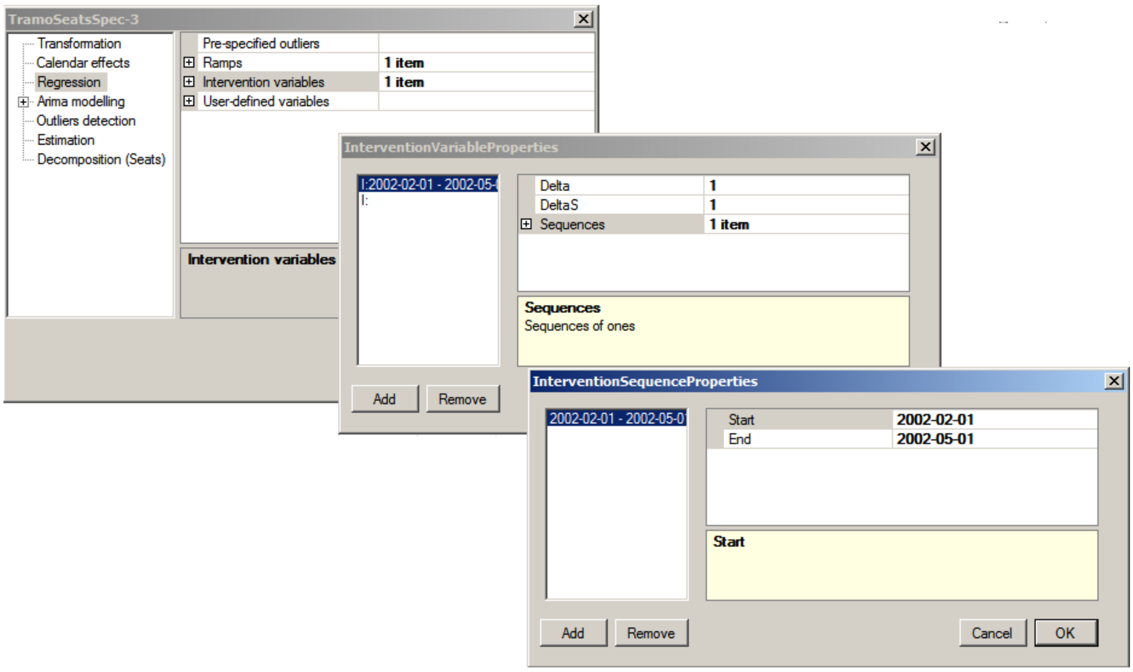



[11] For more details and examples see: MARAVALL, A. (2008).

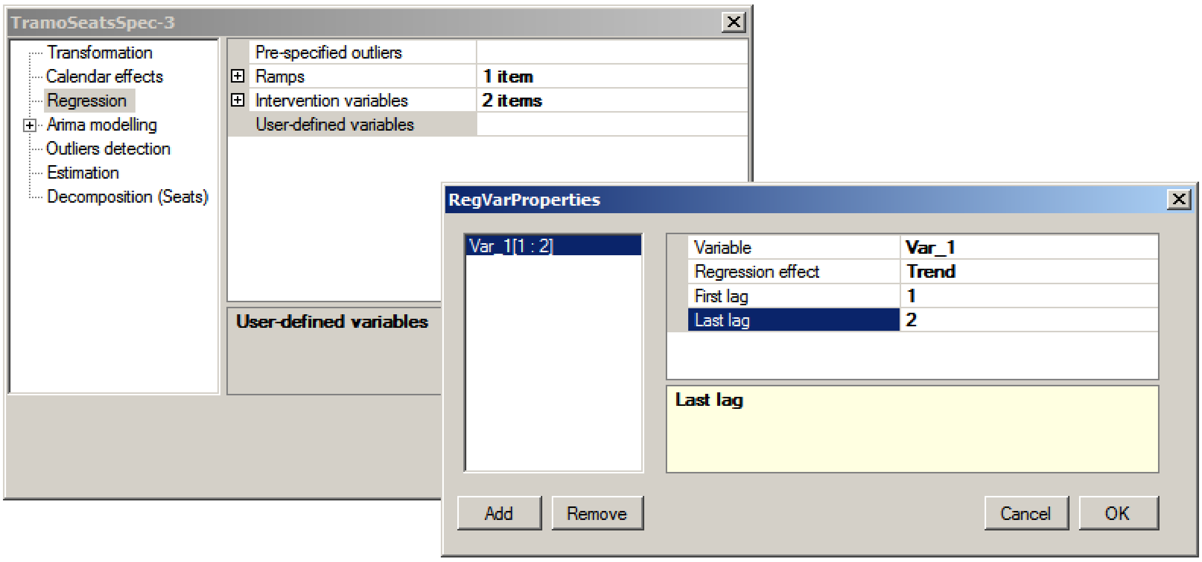
Example: Ramps




Example: Intervention variables

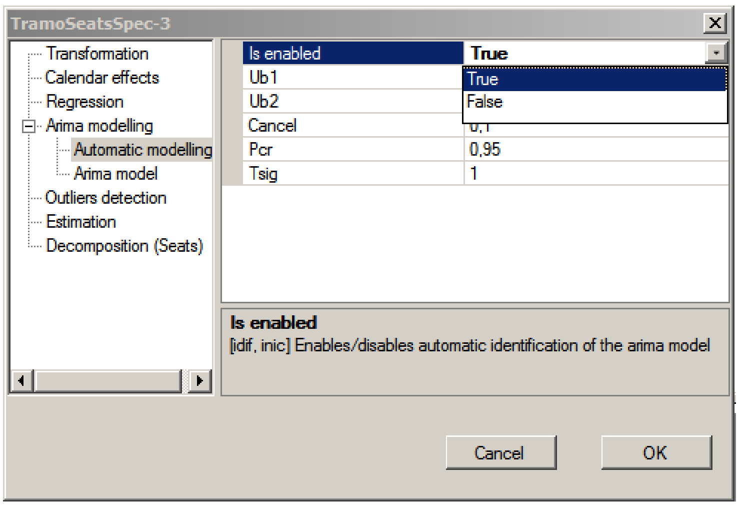


Example: User-defined variables




### 4.2.2.3 Automatic modeling

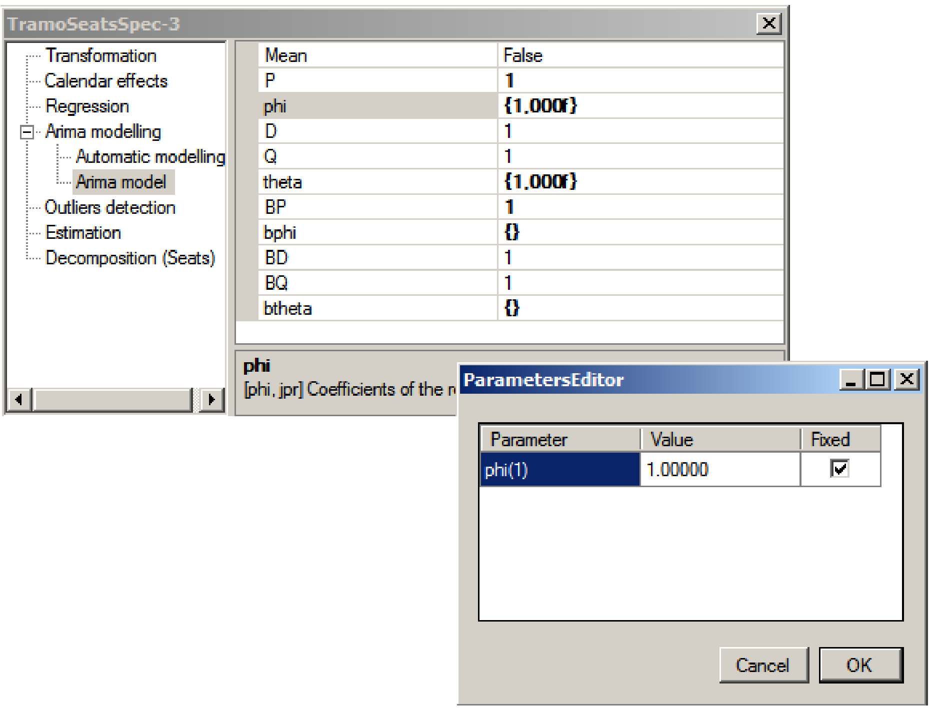
| Item | TramoSeats | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| IsEnabled | *automdl* | | Presence or not of the automdl individual spec |
| Ub1 | *automdl* | *ub1* | Initial unit root limit in the automatic differencing procedure |
| Ub2 | *automdl* | *ub2* | Final unit root limit in the automatic differencing procedure |
| Cancel | *automdl* | *cancel* | Cancelation limit for AR and MA roots |
| Pcr | *automdl* | pcr | Ljung-Box Q statistic limit for the acceptance of a model |
| Tsig | *automdl* | *tsig* | Minimum t for significant mean |



#### *4.2.2.4 Arima*

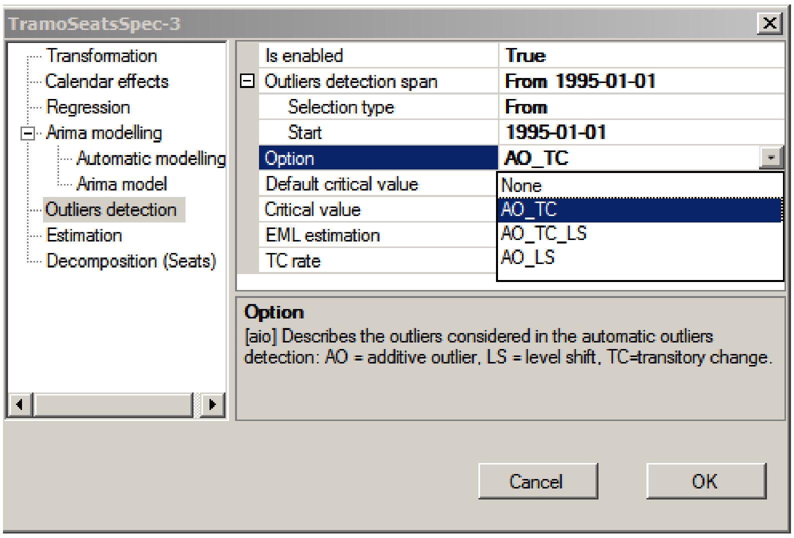
Options included in this section are active only if *IsEnabled* parameter from Automatic modeling section is set to false.

| Item | TramoSeats | | Comments |
|---|---|---|---|
| | **Individual spec** | **Argument** | |
| Mean | regression | variables | It is considered that the mean (a constant term) is part of the Arima model (it highly depends on the chosen model). |
| P, D, Q, BP, BD, BQ | arima | model | Only "Box-Jenkins" SARIMA models (p d q)(bp bd bq) is considered |
| theta, btheta | arima | *[th, jqr]* | Coefficients of the regular (theta) and seasonal (btheta) moving average polynomial |
| phi, bphi | arima | | Coefficients of the regular (phi) and seasonal (bphi) auto-regressive average polynomial. Parameters can be set if P or BP (respectively) is greater than 0. |



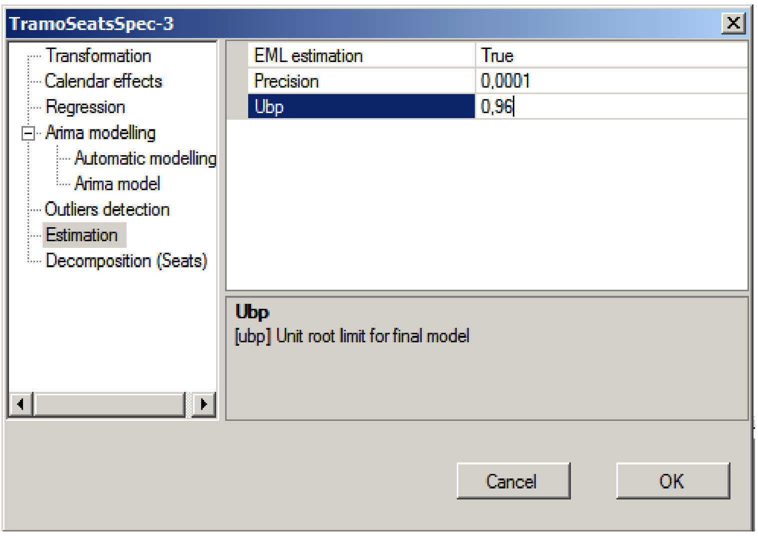
#### 4.2.2.5 Outliers detection

| Item | TramoSeats | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| IsEnabled | *outlier* | | Presence or not of the outlier individual spec |
| Outliers detection span | *outlier* | *span* | Span used for the outlier detection. The span can be computed dynamically on the series (for instance "Excluding last 12 obs"). |
| Option | *outier* | *aio* | Describes the outliers considered in the automatic outliers detection. It is possible to detect all types of outliers, only AO (additive outliers) and TC (transitory change), or only AO and LS. |
| Default critical value | *outlier* | *critical* | When "*Use default critical value*" is false, the procedure uses the critical value mentioned in the specification. Otherwise, the default is used (the first case corresponds to "critical = xxx"; the second corresponds to a spec without the critical argument). It should be noted that it is not possible to define different critical values for different outliers' types. |
| Critical value | *outlier* | *critical* | Critical value used in the outliers' detection procedure. |
| TC rate | *outlier* | *tcrate* | Rate of decay for transitory change outlier regressor. |
| EML estimation | *outlier* | *imvx* | True if exact likelihood estimation method is used, false if fast Hannan-Rissanen method is used. |



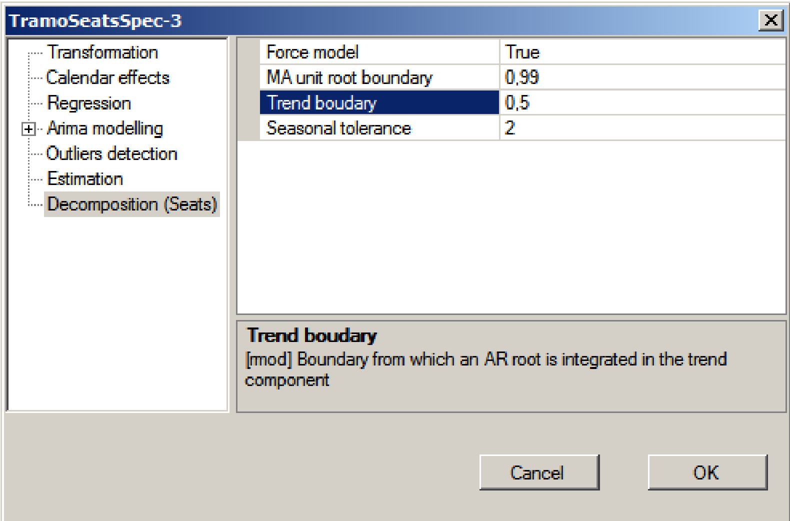
### 4.2.2.6 Estimation

| Item | TramoSeats | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| EML estimaton | | | True if exact likelihood estimation method is used, false if fast Hannan-Rissanen method is used |
| Precision | *estimate* | *tol* | Precision used in the optimization procedure |
| Udp | | *udp* | Unit root limit for final model |




### 4.2.2.7 Decomposition (Seats)

| Item | TramoSeats | | Comments |
|---|---|---|---|
| | Individual spec | Argument | |
| Force model | *seats* | *noadmiss* | When model does not accept an admissible decomposition, force to use an approximation. |
| MA unit root boundary | *seats* | *xl* | When the modules of n estimated root falls in the range (*xl*,1), it is set to 1 if it is in AR; if root is in MA, it is set equal to *xl.* |
| Trend boundary | *seats* | *mod* | Boundary from which an AR root is integrated in the trend component. |
| Seasonal tolerance | *seats* | *epsphi* | Tolerance (measured in degrees) to allocate AR roots into the seasonal component, |




## 4.3 Single processing

Demetra+ offers several ways to define seasonal adjustment of a single time series. A first question, which will determine the best way to proceed, concerns the specification that will be used to start the analysis.

### 4.3.1 Defining a single-processing

The first step to produce a fast seasonal adjustment is to create a processing. The user can take the existing specification or create completely new specification. First category includes pre-defined specifications and specifications previously defined and saved by the user. The second solution is to create the new specification for the needs of seasonal adjustment of particular time series. This can be done when a user wants to use in a frequent way a specification that is not available in the list of the predefined ones (for example if he wants to integrate systematically its own calendar variables or if he want to exclude some kinds of outliers). After creating a new specification it can be added to the user's workspace.

#### *4.3.1.1 Creation of a single processing using existing specification*

Single processing can be launched in two different ways:

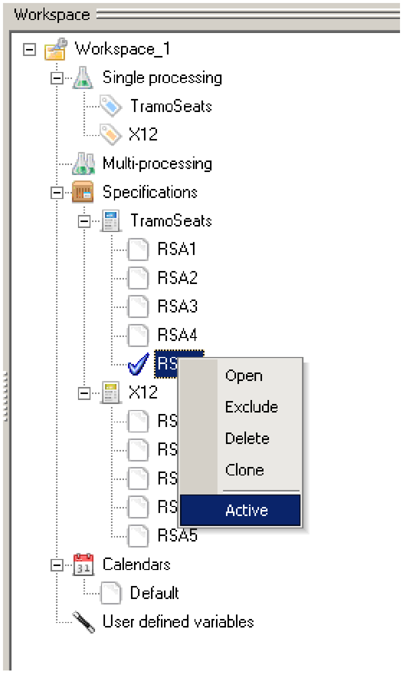
1. ***by activating the specification or drag/drop the specification***

The user could activate the specification from the list displayed in the workspace panel before choosing the series. By default, RSA5c is ticked.

The procedure is as follows:

- Select in the *Workspace tree* the specification you want to activate,

- Open the local menu by means of the right button of the mouse,
- Choose the **Active** option from pop-up menu.

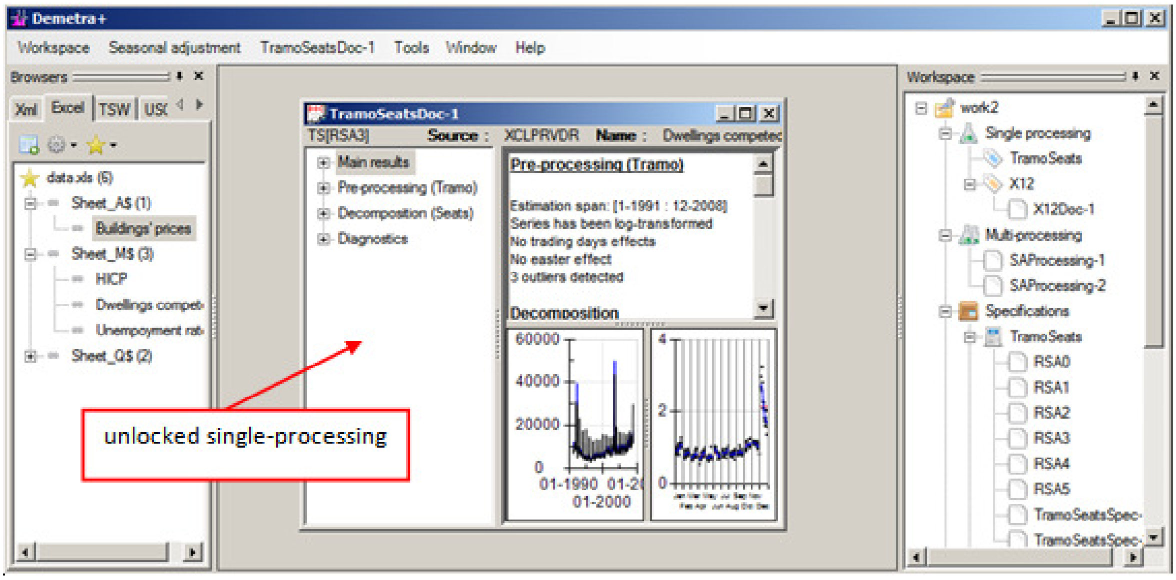



That specification, called active specification, will be used to generate the processing. This specification can be changed at any time.

When the user double-clicks a series in a browser, the software follows the following logic:

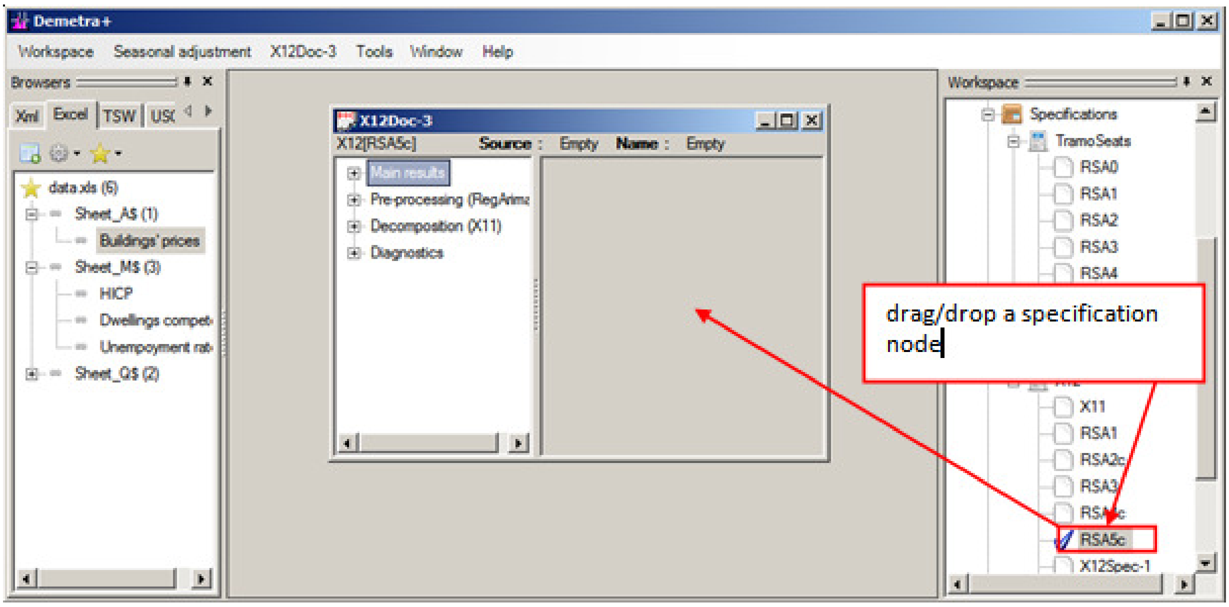
If there is an active specification in the workspace panel, then:

- If some single-processing are open (i.e. single-processing windows have been opened in the central panel), they are updated with the new series.




- If no unlocked single-processing is available, a new one is generated with the active specification.

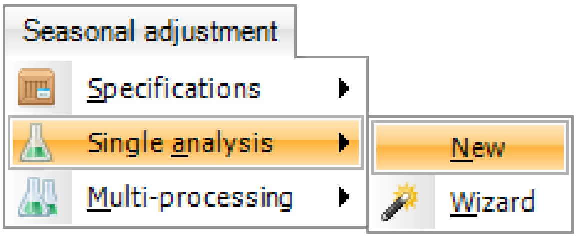
The other option is to drag any specification from the workspace panel and drop it in the central panel of the application. A new single processing window will open automatically.




The data can be imported into specification's window either by a double click on the series of the browsers or by dragging/dropping the series in the left panel of the single processing window.

***2. by the main menu***

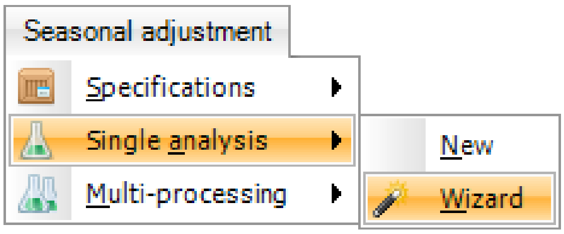
Other method to define single seasonal adjustment is to use the "New" option from the main menu:



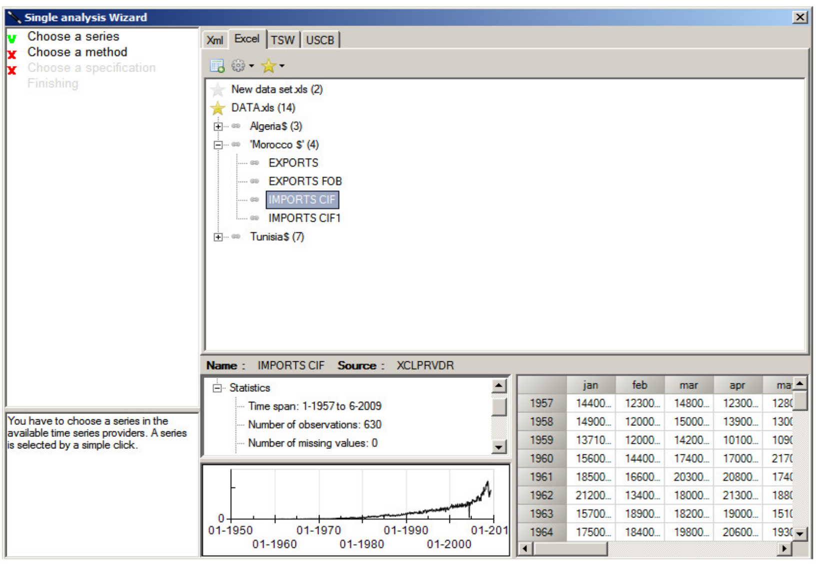

In the next step click "Choose the method"" from the left-hand list and mark TramoSeats or X12. Then, choose the specification from the list of specifications (the contents of the list depends on the method chosen) or define the new one. Demetra+ displays the window with the chosen specification. The last step is to drag the time series from the browser and drop it in this window. The output will be generated instantly.

### *4.3.1.2 Creation of a single processing by defining new specifications*

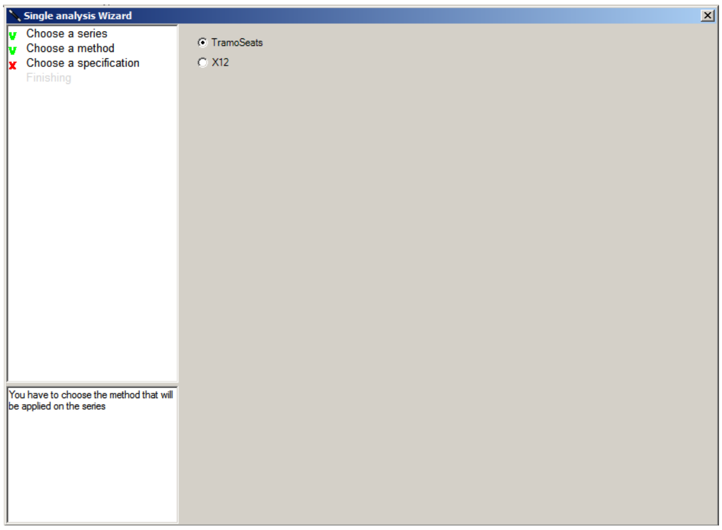
This function is activated from the main menu:



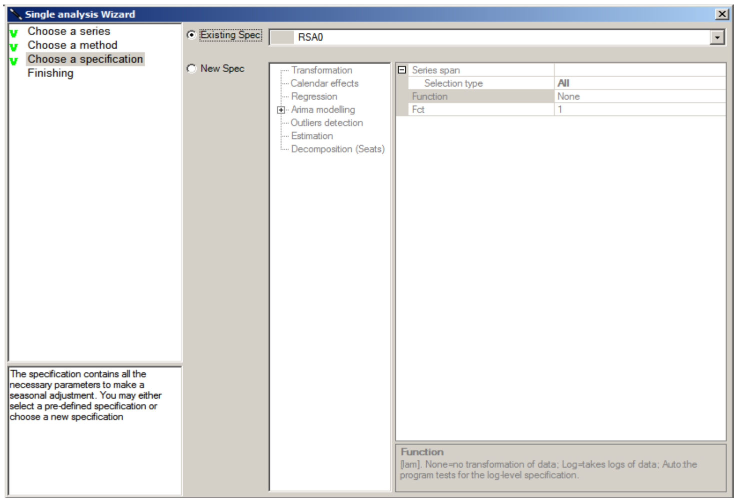

In the first step the user should choose the series he/she wants to analyse, using the browser:



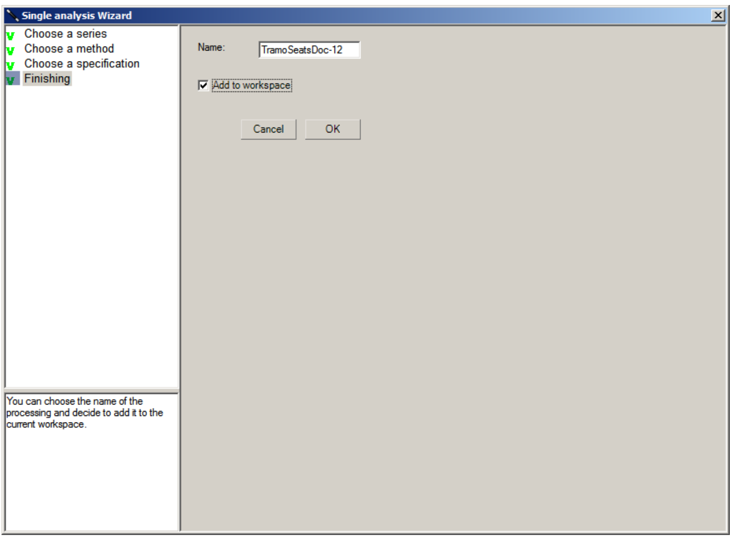

Then the methods could be selected:




After that, the user can choose the specification from the list available in the very top of the window, or create new specification. In the example below the RSA0 specification will be used for seasonal adjustment.



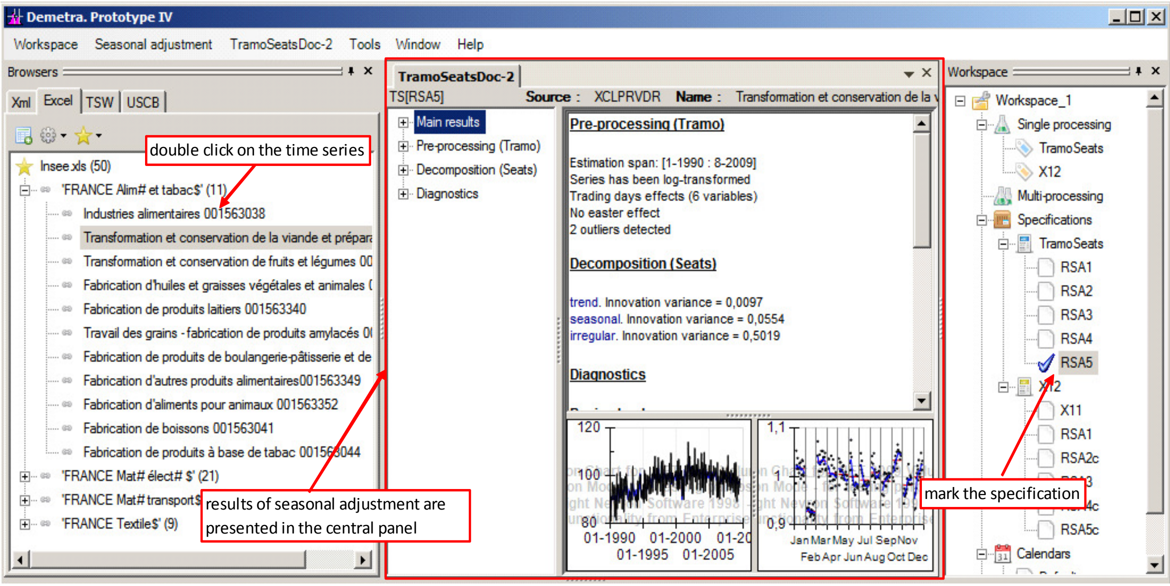

Finally click the "Finishing" item and decide if you want add it into Workspace.



Obviously the user can define the new specification. The specification parameters depend on the method (TramoSeats or X12) chosen in the previous step. For X12 please refer to 4.1. TramoSeats specification is described in 4.2.

### 4.3.2 Seasonal adjustment results – single processing

Specifications correspond to the terminology used in TramoSeats[12] and are described in Annex. Once the active specification is chosen, the user just has to make a double click on the series in the browsers' panel that he wants to adjust. The processing is immediately initiated, with the selected specification and the chosen series.




The user can inspect the different facets of the results through the exploring tree displayed in the left panel of the output window. The results contain many detailed panels. The user can go through them by selecting a node in the navigation tree of the X12 processing. The current specification and the current series are displayed on the top of the window. Demetra+ presents several charts and tables with the results of seasonal adjustment and a set of measures of the quality of seasonal adjustment.

The quality diagnostic implemented in original seasonal adjustment algorithms are different for each SA method. Moreover, their interpretation could be problematic for an unsophisticated user. For this reason, in Demetra+ the qualitative indicator was build-in. Indicator's values are presented in the following table:

12 Description from CAPORELLO, G., and MARAVALL, A. (2004).

*Meaning of the quality indicator*[13]

| ***Value*** | ***Meaning*** |
|---|---|
| Undefined | The quality is undefined because of unprocessed test, meaningless test, failure in the computation of the test, etc. |
| Error | There is an error in the results. The processing should be rejected (for instance, it contains aberrant values or some numerical constraints are not fulfilled |
| Severe | There is no logical error in the results but they should not be accepted for some statistical reasons |
| Bad | The quality of the results is bad, following a specific criterion, but there is no actual error and the results could be used. |
| Uncertain | The result of the test is uncertain |
| Good | The result of the test is good |

Several qualitative indicators can be combined following the basic rules. Given a set of n diagnostics, the sum of the results is:

| ***Sum*** | ***Rules*** |
|---|---|
| Undefined | All diagnostics are Undefined |
| Error | There is at least 1 error |
| Severe | There is at least 1 "severe" diagnostic but no error |
| Bad | No error, no severe diagnostics; the average of the (defined) diagnostics (Bad=1, Uncertain=2, Good=3) is < 1.5 |
| Uncertain | No error, no severe diagnostics; the average of the (defined) diagnostics (Bad=1, Uncertain=2, Good=3) is in [1.5, 2.5[ |
| Good | No error, no severe diagnostics; the average of the (defined) diagnostics (Bad=1, Uncertain=2, Good=3) is ≥ 2.5 |

According to the table, errors and severe diagnostics are absorbent results.

The quality of each diagnostics (except for undefined and error) can be parameterized by the user in Tools->Options->Diagnostic menu.

13 The model also contain a flag "Accepted", which simply means that the statistician decided to accept the results, no matter what are the different diagnostics.

#### 4.3.2.1 X-12-ARIMA

The basic output structure is as follows:

- Main results,
  - Charts,
  - Table,
  - S-I ratio,
- Pre-processing(RegArima),
  - Pre-adjustment series,
  - Arima,
  - Regressors,
  - Residulas,
- Decomposition (X-11),
  - A-Tables,
  - B-Tables,
  - C-Tables,
  - D-Tables,
  - E-tables,
  - Quality measures,
- Diagnostics,
  - Seasonality tests,
  - Spectral analysis,
  - Revisions history,
  - Sliding spans,
  - Model stability.

Detailed description of the seasonal adjustment outcomes is presented below.

#### 4.3.2.1.1 Main results

This section includes basic information about pre-processing and the quality of the outcomes.

**Pre-processing (RegArima)**

Estimation span: [1-1991 : 12-2008]
Series has been log-transformed
No trading days effects
Easter effect detected
3 outliers detected

**Diagnostics**

**summary**
Good

**basic checks**
definition: Good (0,000)
annual totals: Bad (0,063)

**visual spectral analysis**
spectral seas peaks: Good
spectral td peaks: Good

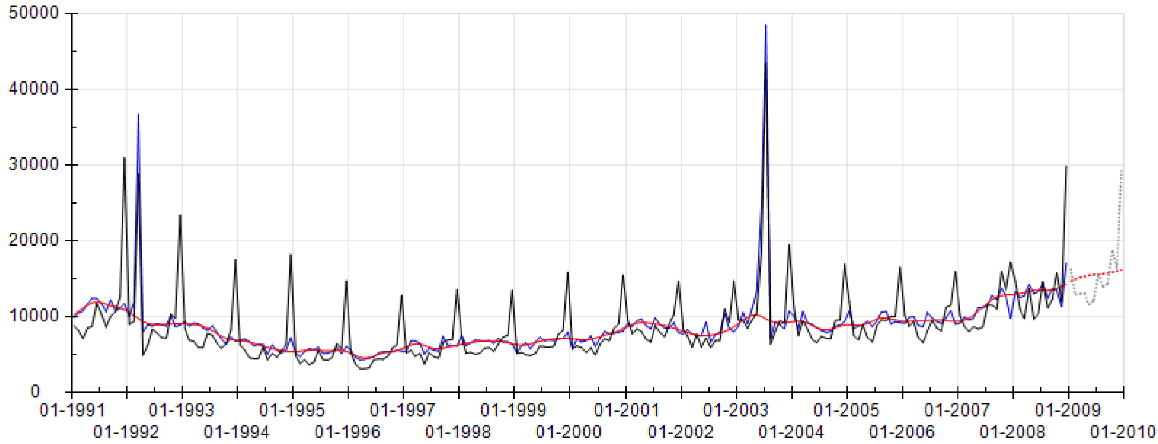
**regarima residuals**
normality: Good (0,740)
independence: Good (0,927)
spectral td peaks: Uncertain (0,098)
spectral seas peaks: Bad (0,003)

**residual seasonality**
on sa: Good (0,606)
on sa (last 3 years): Good (0,926)
on irregular: Good (0,671)

**outliers**
number of outliers: Good (0,014)

**m-statistics**
q: Good (0,437)
q without m2: Good (0,456)

In *Charts* section the top panel presents the original series with forecasts, the final seasonally adjusted series, the final trend with forecasts and the final seasonal component with forecasts. The second panel shows the final irregular component and the final seasonal component with forecasts.



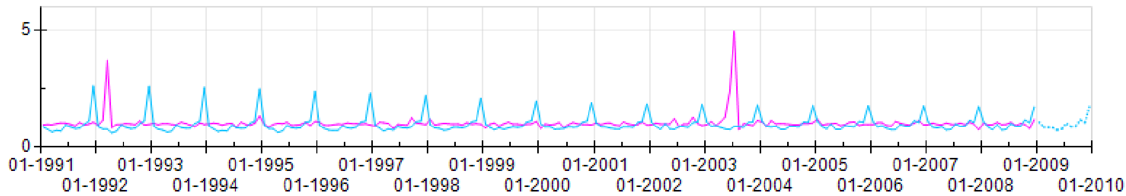
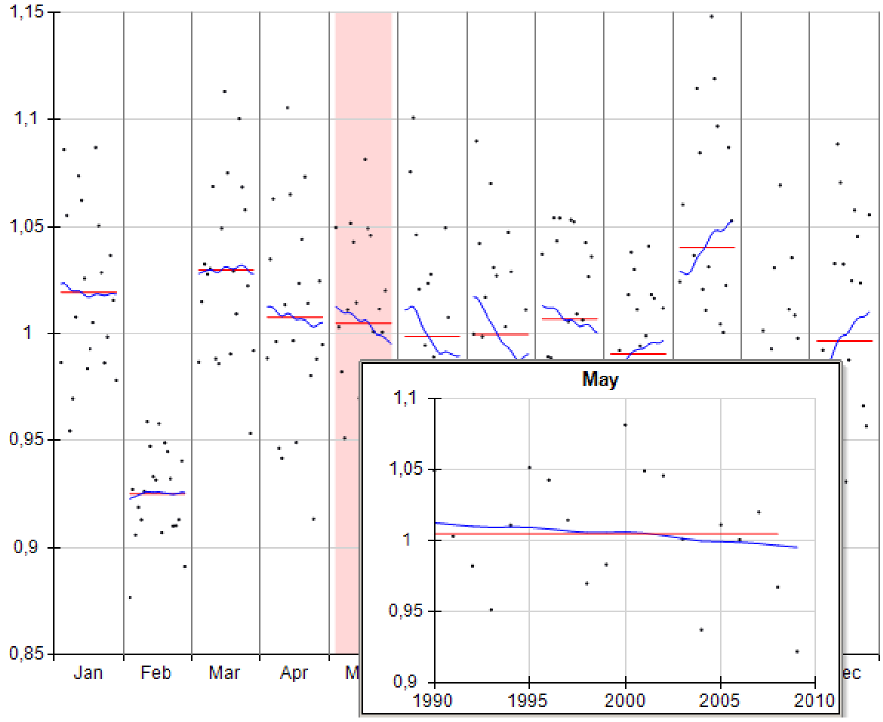




*Table* presents the original series with forecasts and forecast error, the final seasonally adjusted series, the final trend with forecasts, the final seasonal component with forecasts and the final irregular component.

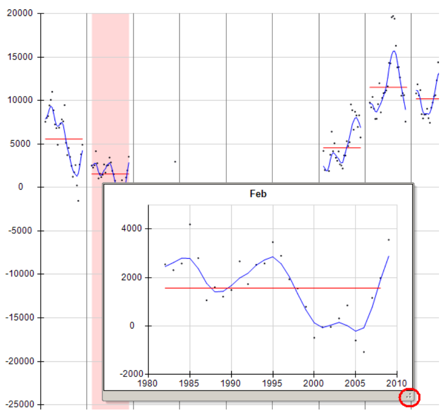
| | Original series | Final seasonally adjusted series | Final trend compone... | Final seasonal... | Final irregula... | C | Fi... | Final trend component (forec |
|---|---|---|---|---|---|---|---|---|
| 1-1991 | 8826 | 9920,93 | 10273,3 | 0,889635 | 0,9657 | | | |
| 2-1991 | 8239 | 10504,2 | 10736,7 | 0,784352 | 0,978344 | | | |
| 3-1991 | 7173 | 10791,4 | 11152 | 0,664698 | 0,967662 | | | |
| 4-1991 | 8586 | 11744 | 11555 | 0,731099 | 1,01635 | | | |
| 5-1991 | 8724 | 12473,7 | 11856,5 | 0,699392 | 1,05206 | | | |
| 6-1991 | 11795 | 12434,4 | 11962 | 0,948577 | 1,0395 | | | |
| 7-1991 | 10358 | 11759,4 | 11880,8 | 0,880827 | 0,989783 | | | |
| 8-1991 | 8618 | 10648,2 | 11687,5 | 0,809336 | 0,911081 | | | |
| 9-1991 | 10104 | 12247,4 | 11483,3 | 0,82499 | 1,06654 | | | |
| 10-19... | 10712 | 10784,9 | 11309,7 | 0,993241 | 0,953596 | | | |
| 11-19... | 12695 | 11165,2 | 11153,9 | 1,13701 | 1,00101 | | | |
| 12-19... | 30960 | 11762,6 | 10945,1 | 2,63208 | 1,07469 | | | |

*S-I ratio* chart presents the final estimation of the seasonal-irregular (SI) component and final seasonal factors for each of the period in time series (months or quarters). Curves represent the final seasonal factors and the straight line represents the mean seasonal factor for each period. The SI ratio presented on the chart (dots) is modified for extreme values (table D9). Final seasonal factors are calculated by applying moving average to the SI ratio from table D9. The results - the final seasonal factors - are displayed in table D10[14].

14 For more details refer to LADIRAY, D., and QUENNEVILLE, B. (1999).




You can enlarge a specific period in the *SI-ratio* chart by clicking in its zone. The details are displayed in a resizable pop-up window (see below).

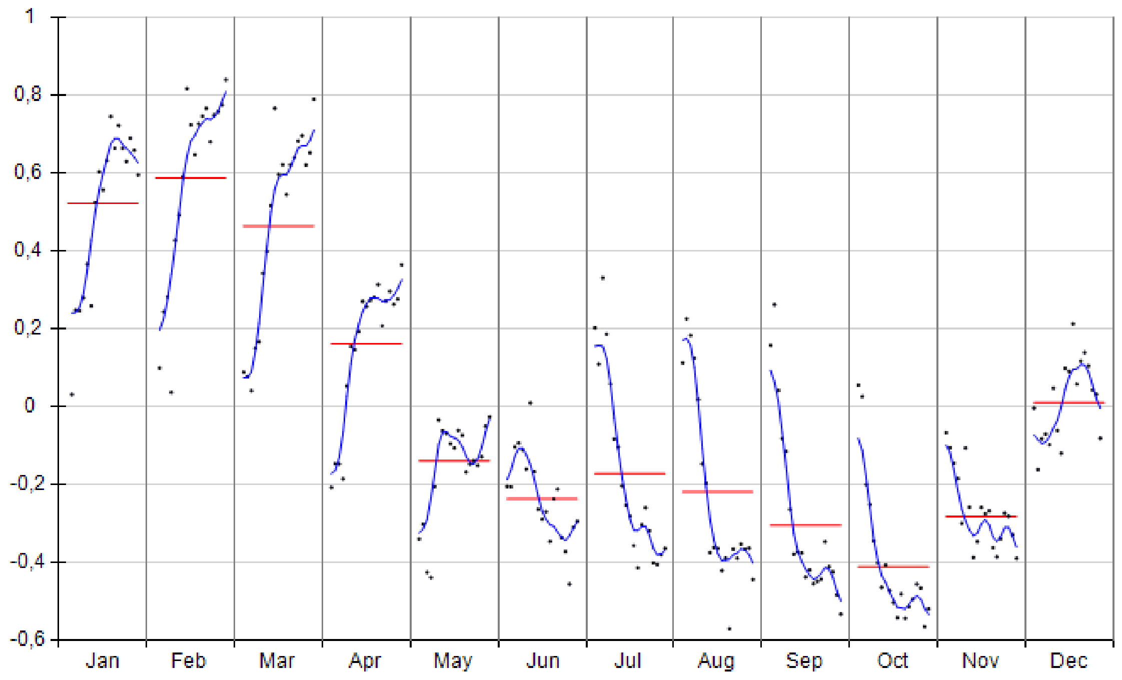


S-I ratio chart is a useful diagnostic tool. This chart is helpful for detecting the presence of seasonal breaks. These would show up as an abrupt changes to the level of the S-I ratios. A seasonal break could distort the estimation of the seasonal component and because of that it should be appropriately modeled[15].

SI-ratio chart also reveals the periods with more statistical variability than other periods. If the SI ratios seem to be very erratic, the seasonal factors will be erratic too. The seasonality is expected to be relatively stable, so in case of high variability of seasonal component the user should choose a longer moving average for its estimation.

15 See: 'Guide To Seasonal Adjustment' (2007).

Changes in seasonality over time are acceptable unless there is a noticeable change from below to above the overall mean. The overall mean is equal to 1 in case of additive model and 0 in case of multiplicative model. The problem is illustrated with the chart below. The S-I ratios for majority of periods are highly unstable. For some of them (e.g. S-I ratios for July, August, September) the effect of seasonality on time series changes from positive to negative. The values of the seasonal component for April indicate that for this period in the beginning of the time series the seasonally adjusted data were higher than raw series while in the end of the period the seasonally adjusted data were smaller than raw series.




#### 4.3.2.1.2 Pre-processing

First part of the pre-processing output includes information about data (estimation span, number of observations actually used in the model, number of parameters in the model, data transformation, correction for leap years) and various information criteria calculated for the model. The  charts below presents exemplary tables taken from different series.

**Data transformation**

Estimation span: [1-2000 : 1-2010]

**Model adequation**

Number of effective observations = 108
Number of estimated parameters = 11

Loglikelihood = -329,7818
Standard error of the regression (ML estimate) = 5,07962
AIC = 681,5635
AICC = 684,3135
BIC = 711,0670
BIC (Tramo definition) = 3,6840
Hannan-Quinn = 693,5261

Next the estimated model parameters are presented. In the example below the Arima model (0,1,0)(0,1,1) was chosen, which means that only one seasonal moving average parameter was calculated. The P-value indicates that the regressor is significant.

**ARIMA model [(0,1,0)(0,1,1)]**

| *Parameter* | *Value* | *Std error* | *T-Stat* | *P-value* |
|---|---|---|---|---|
| BTh(1) | -0,3376 | 0,0882 | -3,83 | 0,0002 |

Using RSA5c specification trading days effect has been detected. It can be noticed from the table below that the regressor for Saturday influences time series in the opposite direction to the other trading days regressors. In spite of the fact that some trading days regressors are insignificant on 5% significance level, the outcome of the join F-test indicates that the trading days regressors are jointly significant.

**Calendar effects**

**Trading days**

| *Parameter* | *Value* | *Std error* | *T-Stat* | *P-value* |
|---|---|---|---|---|
| Monday | 1,56913 | 0,658382 | 2,38 | 0,0191 |
| Tuesday | -2,00054 | 0,649465 | -3,08 | 0,0027 |
| Wednesday | 1,47783 | 0,678829 | 2,18 | 0,0319 |
| Thursday | 0,581356 | 0,659519 | 0,88 | 0,3802 |
| Friday | 0,250291 | 0,632428 | 0,40 | 0,6931 |
| Saturday | -0,829103 | 0,628534 | -1,32 | 0,1902 |
| *Sunday (derived)* | *-1,04896* | *0,659838* | *-1,59* | *0,1151* |

Join F-Test on trading days: F = 4,1350 [P-Value = 0,0010]

If Easter effect was estimated, the following table will be displayed in the output. It is clear, that in the case presented below Easter has a positive, significant effect on the time series (on 1% level).

**Easter**

| Parameter | Value | Std error | T-Stat | P-value |
|---|---|---|---|---|
| easter [15] | 4,45711 | 1,30909 | 3,40 | 0,0009 |

The P-value suggests that leap year effect is insignificant.

**Leap year**

| Parameter | Value | Std error | T-Stat | P-value |
|---|---|---|---|---|
| leap year | 0,331051 | 2,03402 | 0,16 | 0,8710 |

Demetra+ presents also the results of outliers' detection. The table includes the type of outlier, its location, parameter's value and significance.

**Detected outliers**

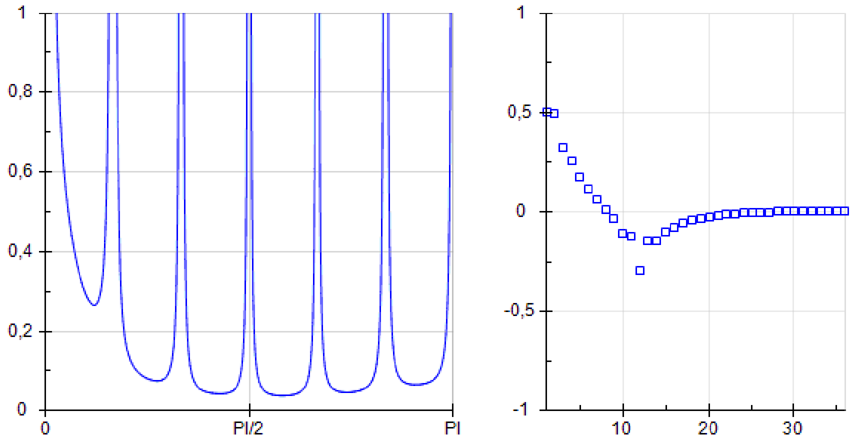
| Parameter | Value | Std error | T-Stat | P-value |
|---|---|---|---|---|
| AO[4-2002] | 18,7046 | 3,41647 | 5,47 | 0,0000 |
| LS[1-2007] | 38,5629 | 4,60194 | 8,38 | 0,0000 |

**Pre-adjustment series**

Table presented in this section contains series estimated by Reg-ARIMA part. It includes interpolated series, series adjusted for calendar effects, deterministic component, calendar effects, trading days effect, outliers effect on irregular component, total outliers effect, total regression effect.

**Arima**

This section demonstrates theoretical spectrum of the stationary and non-stationary model and autocorrelation function of the stationary part of the model.




**Polynomials**

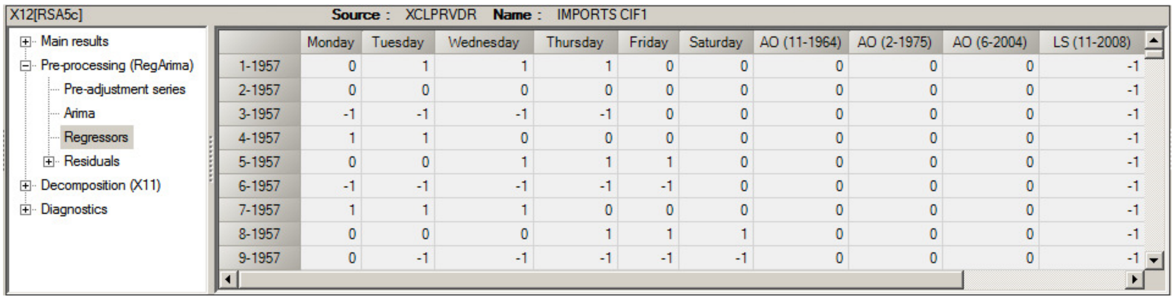
regular AR: 1 - 0,34299 B - 0,32833 B^2
seasonal AR: 1
regular MA: 1
seasonal MA: 1 - 0,35933 S

**Frequency of the regular AR roots**

3,14159265358979

## Regressors

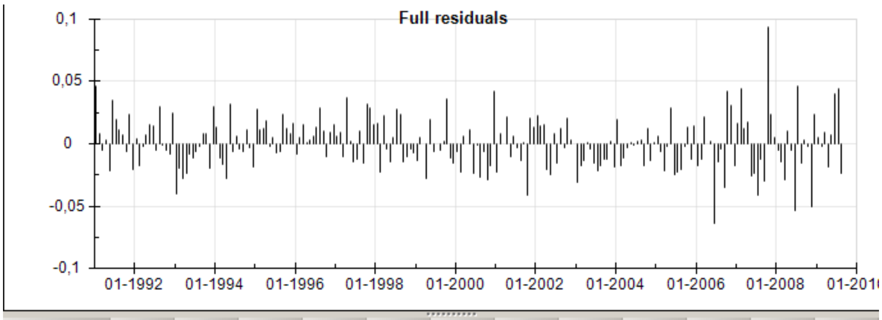
This section presents all regressors used in RegArima part, including trading days variables, leap year effect, outliers, Easter effect, ramps, intervention variables, user-defined variables.




| | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | AO (11-1964) | AO (2-1975) | AO (6-2004) | LS (11-2008) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1-1957 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | -1 |
| 2-1957 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 |
| 3-1957 | -1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | -1 |
| 4-1957 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 |
| 5-1957 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | -1 |
| 6-1957 | -1 | -1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | -1 |
| 7-1957 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | -1 |
| 8-1957 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | -1 |
| 9-1957 | 0 | -1 | -1 | -1 | -1 | -1 | 0 | 0 | 0 | -1 |

## Residuals

The way in which Demetra+ calculates the residuals is presented in Annex. Residuals from the model are presented in the graph and in the table.




| | jan | feb | mar | apr | may | jun | jul | aug | sep | oct | nov | dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1991 | 0,046 | 0,009 | -0,005 | 0,004 | -0,022 | 0,035 | 0,019 | 0,012 | 0,007 | -0,006 | 0,024 | -0,020 |
| 1992 | 0,004 | -0,018 | -0,002 | 0,007 | 0,016 | 0,014 | -0,005 | 0,030 | -0,001 | -0,005 | -0,008 | 0,024 |
| 1993 | -0,040 | -0,020 | -0,028 | -0,024 | -0,008 | -0,011 | -0,006 | -0,002 | 0,008 | 0,008 | -0,020 | 0,029 |
| 1994 | 0,014 | -0,012 | -0,016 | -0,028 | 0,031 | -0,006 | 0,006 | -0,005 | -0,007 | 0,012 | -0,004 | -0,019 |
| 1995 | 0,028 | 0,011 | 0,012 | 0,018 | -0,002 | 0,005 | -0,007 | -0,006 | 0,023 | 0,012 | 0,009 | 0,016 |
| 1996 | -0,009 | 0,005 | 0,015 | 0,002 | 0,003 | 0,006 | 0,014 | 0,029 | 0,011 | -0,010 | 0,009 | 0,015 |
| 1997 | 0,007 | 0,009 | -0,010 | 0,037 | 0,002 | -0,014 | -0,012 | 0,011 | -0,016 | 0,031 | 0,029 | 0,015 |
| 1998 | 0,017 | -0,023 | 0,022 | -0,004 | -0,015 | 0,005 | 0,028 | 0,024 | -0,015 | -0,011 | -0,004 | -0,007 |
| 1999 | -0,013 | 0,005 | 0,000 | -0,028 | 0,019 | -0,006 | 0,000 | -0,005 | 0,002 | 0,036 | -0,012 | -0,016 |
| 2000 | -0,006 | -0,022 | 0,006 | 0,000 | 0,012 | -0,024 | -0,001 | -0,027 | -0,007 | -0,029 | -0,018 | 0,042 |
| 2001 | -0,023 | 0,008 | 0,000 | 0,021 | -0,011 | 0,006 | -0,003 | -0,013 | 0,001 | -0,042 | 0,021 | 0,013 |

Analysis of the residuals consists of several tests, which are described in the Annex. Summary statistics are presented in the following tables:

1. Normality of the residuals

| | P-value |
|---|---|
| Mean | 0,2032 |
| Skewness | 0,2986 |
| Kurtosis | 0,6680 |
| Normality | 0,4859 |

2. Independence of the residuals

| | P-value |
|---|---|
| Ljung-Box(24) | 0,5164 |
| Box-Pierce(24) | 0,5781 |
| Ljung-Box on seasonality(3) | 0,4087 |
| Box-Pierce on seasonality(3) | 0,4442 |

3. Randomness of the residuals

| | P-value |
|---|---|
| Runs around the mean: number | 0,8312 |
| Runs around the mean: length | 1,0000 |
| Up and Down runs: number | 0,7453 |
| Up and Down runs: length | 1,0000 |

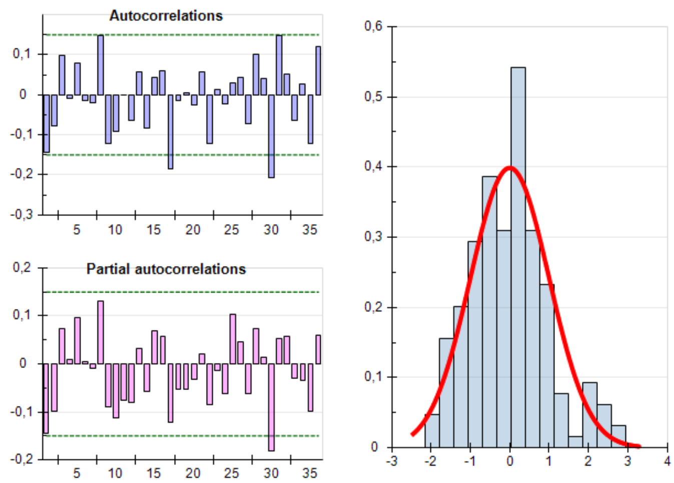
4. Linearity of the residuals

| | P-value |
|---|---|
| Ljung-Box on squared residuals(24) | 0,0028 |
| Box-Pierce on squared residuals(24) | 0,0054 |

For tests one to three the null hypothesis was accepted (p-values higher than 5%). It means that it can be assumed that residuals are independent and random. They are approximately normally distributed.

The p-value marked in red indicates that null hypothesis was rejected. There is an evidence of autocorrelation in residuals. A linear structure is left in the residuals.

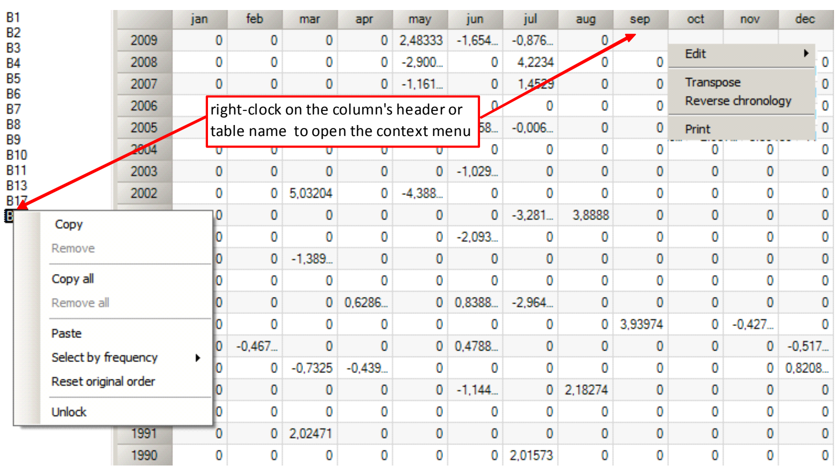
Demetra+ also presents residuals' distribution. In this section autocorrelation and partial autocorrelation functions are presented.




#### 4.3.2.1.3 Decomposition

**Tables**

In this section all important tables from X-11 procedure are available. The view of B-tables is presented below:




The detailed tables' list can be found in the Annex.

**Quality measures**

This section presents the seasonal and trend moving filters used to estimate the seasonal factors and the final trend-cycle. Demetra+ selects the filters automatically, taking into account the global moving seasonality ratio, which is computed on preliminary estimates of the irregular component and of the seasonal.

The M-statistics are used to judge the quality of seasonal adjustment[16]. These statistics vary between 0 and 3 but only values smaller than 1 are acceptable. M1 measures the contribution of the irregular component to the total variance. M2, which is very similar to M1, is calculated on the basis of the contribution of the irregular component to the stationary portion of the variance. Statistic M3 compares the irregular to the trend-cycle taken from preliminary estimate of the seasonally adjusted series, as if this ratio is too large, it is difficult to separate the two components. Statistic M4 tests the randomness of irregular component. Statistic M5 statistic is used to compare the significance of changes in trend with that in irregular. Statistic M6 checks the I/S (irregular/seasonal component ratio) as if annual changes in the irregular component are too small in relation to the annual changes in the seasonal component the $3\times5$ seasonal filter used for estimation of the seasonal component is not flexible enough to follow the seasonal movement. It should be underlined that statistic M6 is calculated only if this filter has been used in the model. Statistic M7 is the combined test for the presence of identifiable seasonality. The test compares the relative contribution of stable and moving seasonality. Statistics M8 to M11 measures if the movement due to short-term quasi-random variations and movement due to long term changes are not changing too much over the years. If the changes are too strong then the seasonal factors could be erroneous.

Q-statistic is a composite indicator calculated from M-statistics. Q without M2 (also called Q2) is the Q without the M2 statistics.

$$Q = \frac{10M1 + 11M2 + 10M3 + 8M4 + 11M5 + 10M6 + 18M7 + 7M8 + 7M9 + 4M10 + 4M11}{100}$$

Otherwise its weight in q is zero. If time series does not cover at least 6 years statistics M8, M9, M10 and M11 cannot be calculated. In this case the Q statistics is calculated as:

$$Q = \frac{14M1 + 15M2 + 10M3 + 8M4 + 11M5 + 10M6 + 32M7 + 0M8 + 0M9 + 0M10 + 0M11}{100}$$

The model has a satisfactory quality if Q statistic is less than 1.

*Results of the test*

| ***Treshold*** | ***Diagnostic*** |
|---|---|
| ≥2 | Severe |
| [1,2[ | Bad |
| <1 | Good |

[16] For the definition of the M statistics refer to Ladiray D., and Quenneville B. (1999).

***Summary and Quality measures***

**Final fiters**

Trend filter: 9-term Henderson moving average
Seasonal filter: 3 x 5 moving average

**Relative contribution of the components to the stationary portion of the variance in the original series**

| I | C | S | P | TD&H | Total |
|---|---|---|---|---|---|
| 0,69 | 65,75 | 27,11 | 2,74 | 4,95 | 101,24 |

**Monitoring and Quality Assessment Statistics**

| Measure | Value |
|---|---|
| M1 | 0,099 |
| M2 | 0,071 |
| M3 | 0,000 |
| M4 | 0,000 |
| M5 | 0,201 |
| M6 | 0,341 |
| M7 | 0,209 |
| M8 | 0,496 |
| M9 | 0,296 |
| M10 | 0,584 |
| M11 | 0,000 |

Summary

| | |
|---|---|
| Q | **0,212** |
| Q without M2 | **0,229** |

#### 4.3.2.1.4 Diagnostics

The *Diagnostic* panel contains detailed information on the seasonal adjustment process.

**summary**
Good

**basic checks**
definition: Good (0,000)
annual totals: Bad (0,063)

**visual spectral analysis**
spectral seas peaks: Good
spectral td peaks: Good

**regarima residuals**
normality: Good (0,740)
independence: Good (0,927)
spectral td peaks: Uncertain (0,098)
spectral seas peaks: Bad (0,003)

**residual seasonality**
on sa: Good (0,606)
on sa (last 3 years): Good (0,926)
on irregular: Good (0,671)

**outliers**
number of outliers: Good (0,014)

**m-statistics**
q: Good (0,437)
q without m2: Good (0,456)

- **Basic checks**

This section offers a set of quality diagnostic.

  - **Definition**

This test is inspecting some basic relationships between different components of the time series. The following components are defined[17]:

| ***Name*** | ***Code*** | ***Definition*** |
|---|---|---|
| Y | y(_f) | Original series |
| Yc | yc(_f) | Interpolated series (= Y with missing values relaced by their estimates) |
| T | tl(_f) | Trend (without regression effects) |
| S | sl(_f) | Seasonal (without regression effects) |
| I | il | Irregular (without regression effects) |
| SA | sal | Seasonally adjusted series(without regression effects) |
| SI | | SI ratio |
| TDE | td(_f) | Trading days effects |
| MHE | mh(_f) | Moving holidays effects |
| EE | | Easter effects |
| RMDE | | Ramadan effects |
| OMHE | | Other moving holidays effects |
| CAL | cal(_f) | Calendar effects |
| OTOT<br>O(cmp=T,S,I) | out<br>out_t, out_s, out_i | Outliers effects |
| REGTOT<br>REG(cmp=Y, SA, T,S, I) | reg(_f)<br>reg_y(_f),<br>reg_sa(_f)<br>reg_t(_t),<br>reg_s(_f), reg_i(_f) | Other regression effects |
| DET(cmp=T,S, I,Y) | det(_f)<br>det_y(_f),<br>det_sa(_f)<br>det_t(_f),<br>det_s(_f), det_i(_f) | Deterministic effects |
| C(cmp=T,S,I,SA) | t(_f), s(_f), i, sa | Components, including deterministic effects |
| Ycal | ycal | Calendar corrected series |
| Yl | yl | Linearized series |

17 The names mentioned in the document appear in the graphical interface of Demetra+. The corresponding codes are used in the csv output. For compatibility issues with previous versions, they have not been aligned on the names. For some series, it is possible to generate the forecasts (computed on 1 year); the corresponding code is defined by adding the "_f" suffix (for example, y becomes y_f).

For those components in additive case the following relationships should be true:

$$MHE = EE + RMDE + OMHE \quad (1)$$

$$CAL = TDE + MHE \quad (2)$$

$$OTOT = OT + OS + OI \quad (3)$$

$$REGTOT = REGT + REGS + REGI + REGY \quad (4)$$

$$REGSA = REGT + REGI \quad (4)$$

$$DET = CAL + OTOT + REGTOT \quad (5)$$

$$CT = T + OT + REGT \quad (6)$$

$$CS = S + CAL + OS + REGS \quad (7)$$

$$CI = I + OI + REGI \quad (8)$$

$$CSA = Y_c - CS = CT + CI + REGY \quad (9)$$

$$Y_c = CT + CS + CI + REGY = T + S + I + DET \quad (10)$$

$$Y_l = Y_c - DET = T + S + I \quad (11)$$

$$SA = Y_l - S = T + I \quad (12)$$

$$S_l = Y_l - T = S + I \quad (13)$$

The multiplicative model is obtained in the same way by replacing the operations "+" and "-" by "*" and "/" respectively.

A first test in *Basic diagnostic* verifies that all the definition constraints are well respected. The maximum of the absolute differences is computed for the different equations and related to the Euclidean norm of the initial series (Q).

*Results of the test*

| ***Q*** | ***Diagnostic*** |
|---|---|
| > 0.000001 | **Error** |
| <= 0.000001 | Good |

- **Annual totals**

The test compares the annuals totals of the original series and those of the seasonally adjusted series. The maximum of their absolute differences is computed and related to the Euclidean norm of the initial series.

*Results of the test*

| ***Q*** | ***Diagnostic*** |
|---|---|
| > 0.5 | Error |
| ]0.1, 0.5] | Severe |
| ]0.05, 0.1] | Bad |
| ]0.01, 0.05] | Uncertain |
| <=0.01 | Good |

- **Visual spectral analysis**

Demetra+ identifies spectral peaks in seasonal ad trading days components using empiric criterion of “visual significance”. For more information see the Annex.

- **RegArima Residuals diagnostics**

Several tests are computed on the residuals of the RegArima model. The exact definition of what we mean by "residuals" should be clarified. Indeed, X12 and Tramo are based on different estimation procedures of the likelihood of the RegArima models, which lead to different definitions of the residuals. Demetra+ takes still another way (similar to the solution developed in Stamp, for instance). The Annex describes those solutions.

In most cases, the different sets of residuals yield slightly different diagnostics. However, their global messages are nearly always very similar[18].

- **Normality test**

The joint normality test (which combines skewness and kurtosis tests) is the Doornik-Hansen test (see appendix 3), which is distributed as a $\chi^2$ (2).

*Results of the test*

| ***Pr(*** $\chi^2$ **(2) *>val)*** | ***Diagnostic*** |
|---|---|
| <0.01 | Bad |
| [0.01, 0.1[ | Uncertain |
| ≥0.1 | Good |

- **Independence test**

18 In future versions of Demetra+, it will be possible to choose the definition of the residuals that must be used in the tests and displayed in the graphical interface. Obviously, the choice is more a question for purists.

The independence test is the Ljung-Box test (see Annex), which is distributed as $\chi^2(k-np)$,

where $k$ depends on the frequency of the series (24 for monthly series, 8 for quarterly series, $4*freq$ for other frequencies, where $freq$ is a frequency of the time series) and $np$ is the number of hyper-parameters of the model (number of parameters in the Arima model).

*Results of the test*

| ***Pr(*** $\chi^2(k-np)$ ***>val)*** | ***Diagnostic*** |
|---|---|
| <0.01 | Bad |
| [0.01, 0.1[ | Uncertain |
| ≥0.1 | Good |

- **Spectral tests**

Demetra+ testing the presence of the trading days and seasonal peaks in the residuals. For this purpose the tests based on the periodogram of the residuals, for the trading days frequencies and for the seasonal frequencies are implemented.

The periodogram is computed at the so-called Fourier frequencies, which present good statistical properties. Under the hypothesis of Gaussian white noise of the residual, it is possible to derive simple tests on the periodogram, around specific (groups of) frequencies. The exact definition and the used test are described in the Annex.

*Results of the test*

| ***P(stat>val)*** | ***Diagnostic*** |
|---|---|
| <0.001 | Severe |
| [0.001, 0.01[ | Bad |
| [0.01, 0.1[ | Uncertain |
| ≥0.1 | Good |

- **Residual seasonality diagnostics**

The residual seasonality diagnostics correspond to the tests developed in X12.

The F-Test on stable seasonality (see Annex) is computed on the differences of the seasonally adjusted series (component CSA, see above) and on the irregular component (CI, see above).

In order to extract the trend from the monthly time series a first order difference of lag three is applied (a first order difference of lag one in the other cases)[19]. For the seasonally adjusted series, one test is computed on the complete time span and another one on the last 3 years.

[19] DAGUM, E. B. (1987).

*Results of the test*

| ***Pr(F>val)*** | ***Diagnostic*** |
|---|---|
| <0.01 | Severe |
| [0.01, 0.05[ | Bad |
| [0.05, 0.1[ | Uncertain |
| ≥0.1 | Good |

- **Number of outliers**

High number of outliers indicates that there is a problem related to a weak stability of the process or the reliability of the data is low. If the high number of outliers was detected (above 3%, according to the table), the chosen ARIMA model cannot fit all of the observations.

*Results of the test*

| ***Treshold*** | ***Diagnostic*** |
|---|---|
| ≥0.1 | Severe |
| [0.05, 0.1[ | Bad |
| [0.03, 0.05[ | Uncertain |
| <0.3 | Good |

- **M-statistics**

For the test results refer to 4.3.2.1.3.

### *Seasonality tests*

The diagnostic section includes the set of seasonality test useful for checking the presence of seasonality in time series. Those tests are described in Annex.

The seasonal component includes the intra-year variation that is repeated constantly (stable seasonality) or evolving from year to year (moving seasonality). To determine if stable seasonality if present in a series Demetra+ computes Friedman test using the seasons (months or quarters) as the factor on the preliminary estimation of the unmodified SI component.

**Friedman test**

Friedman statistic = 73,4321
Distribution: F-stat with 11 degrees of freedom in the numerator and 154 degrees of freedom in the denominator
P-Value: 0,0000
Stable seasonality present at the 1 per cent level

A large test statistics and small significance level indicates that a significant amount of variation in the SI-ratios is due to months (or quarters, respectively), which in turn is evidence of seasonality.

If the p-Value is lower than 0.1% the null hypothesis of no seasonal effect is rejected. Conversely, a small F and large significance level (close to 1.0) is evidence that variation due to month or quarter could be due random error and the null hypothesis of no month/quarter effect is not rejected[20].

In the example above p-Value is 0.0000, so the null hypothesis is rejected and it could be assumed that significant seasonality is present.

The second test for stable seasonality provided by Demetra+ is Kruskal-Wallis test.

**Kruskall-Wallis test**

Kruskall-Wallis statistic = 162,5477
Distribution: Chi2(11)
P-Value: 0,0000
Stable seasonality present at the 1 per cent level

The test's outcome (stable seasonality present) has confirmed the result from Friedman test.

The test for the presence of seasonality assuming stability uses the following decomposition of the variance: $S^2 = S_A^2 + S_R^2$ where:

$S^2 = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_{ij} - \bar{X})^2$ - the total sum of squares,

$S_A^2 = \sum_{j=1}^{k} n_j (\bar{X}_{\bullet j} - \bar{X}_{\bullet\bullet})^2$ - variance of the averages, due to seasonality,

$S_R^2 = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_{\bullet j})^2$ - the residual sum of squares.

Explanation of the test and symbols is included in Annex.

The test statistics is calculated as:

$$F_S = \frac{\frac{S_A^2}{k-1}}{\frac{S_R^2}{n-k}} \sim F(k-1, n-k)$$

where $k-1$ and $n-k$ are degrees of freedom.

20 http://support.sas.com/onlinedoc/913

The example is shown below:

**Test for the presence of seasonality assuming stability**

| | Sum of squares | degrees of freedom | Mean square |
|---|---|---|---|
| Between months | 0,3532 | 11 | 0,0321 |
| Residual | 0,0544 | 179 | 0,0003 |
| Total | 0,4075 | 190 | 0,0021 |

Value: 105,6788
Distribution: F-stat with 11 degrees of freedom in the numerator and 179 degrees of freedom in the denominator
P-Value: 0,0000
Seasonality present at the 1 per cent level

The test statistic was calculated in the following way:

$$F_S = \frac{\frac{0,3532}{12-1}}{\frac{0,0544}{191-12}} \sim F(11,179)$$

The p-Value is $0,0000$ so the null hypothesis is rejected and it could be assumed that the seasonality in time series is significant.

**Evolutive seasonality test**

| | Sum of squares | Degrees of freedom | Mean square |
|---|---|---|---|
| Between years | 0,0021 | 14 | 0,0001 |
| Error | 0,0305 | 154 | 0,0002 |

Value: 0,7419
Distribution: F-stat with 14 degrees of freedom in the numerator and 154 degrees of freedom in the denominator
P-Value: 0,7296
No evidence of moving seasonality at the 20 per cent level

**Combined seasonality test**

Identifiable seasonality present

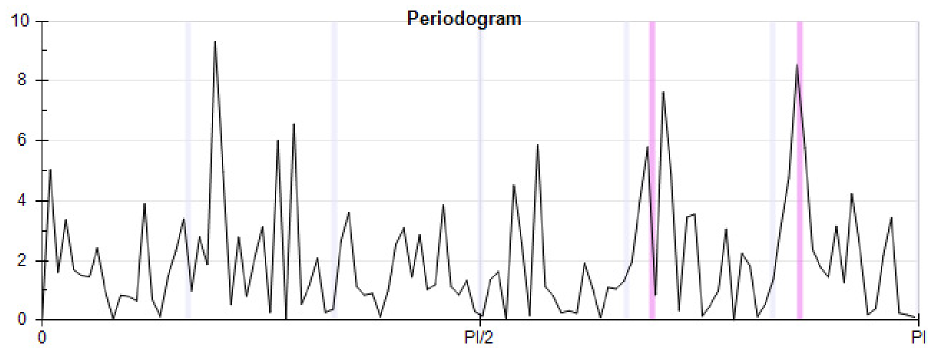
**Residual seasonality test**

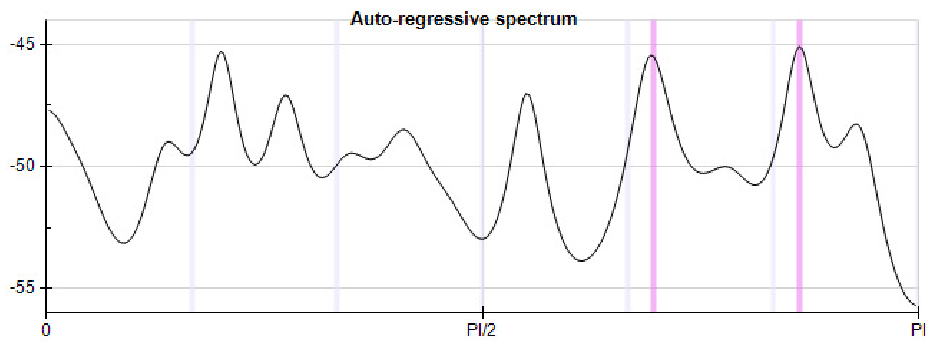
No evidence of residual seasonality in the entire series at the 10 per cent level: F=0,2132
No evidence of residual seasonality in the last 3 years at the 10 per cent level: F=0,3454

### *Spectral analysis*

Demetra+ provides spectral plots to alert the user to the presence of remaining seasonal and trading day effects. The graphics are available for residuals, irregular component and seasonally adjusted time series. In order to compare the results with spectral analysis for raw time series, the user should create the relevant graph for raw time series from Tools menu (see: 3.2).

Two spectrum estimators are implemented: periodogram and auto-regressive spectrum[21]. Seasonal frequencies are marked as grey, vertical lines, while violet lines correspond to trading-days frequencies. The X-axis shows the different frequencies. The periodicity of phenomenon at frequency $f$ is $\frac{2\pi}{f}$. It means that for monthly time series the seasonal frequencies are: $\frac{\pi}{6},\frac{\pi}{3},\frac{\pi}{2},\frac{2\pi}{3},\frac{5\pi}{6}$ (which are equivalent to $\frac{1}{12},\frac{2}{12},\ldots$ cycles per month, i.e. in the case of a monthly series, the frequency $\frac{\pi}{3}$ corresponds to a periodicity of 6 months.). The trading days frequencies are described in Annex. Peak at the zero frequency corresponds to the trend component of the series.




At seasonal and trading days frequencies, a peak in model residuals indicates the need for a better fitting model. In particular, peaks at the seasonal frequencies are caused by inadequate filters chosen for decomposition. Peaks at the trading days frequencies could occur due to inappropriate regression variables used in the model or the significant change of the calendar effect because the calendar effect cannot be modeled by fixed regression effect on the whole time series span.

A peak in the spectrum from the seasonally adjusted series or irregulars reveals inadequacy of the seasonal adjustment filters for the time interval used for spectrum estimation. In this case different model specification or data span length should be considered.

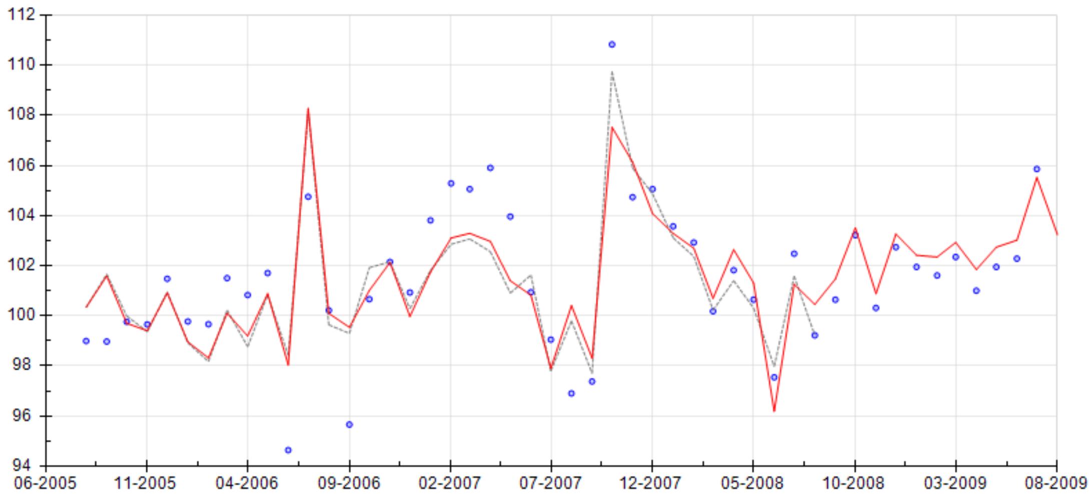
21 The theoretical motivation for the choice of spectral estimator is provided by SOKUP, R. J., and FINDLEY, D. F. (1999).

### *Revision histories*

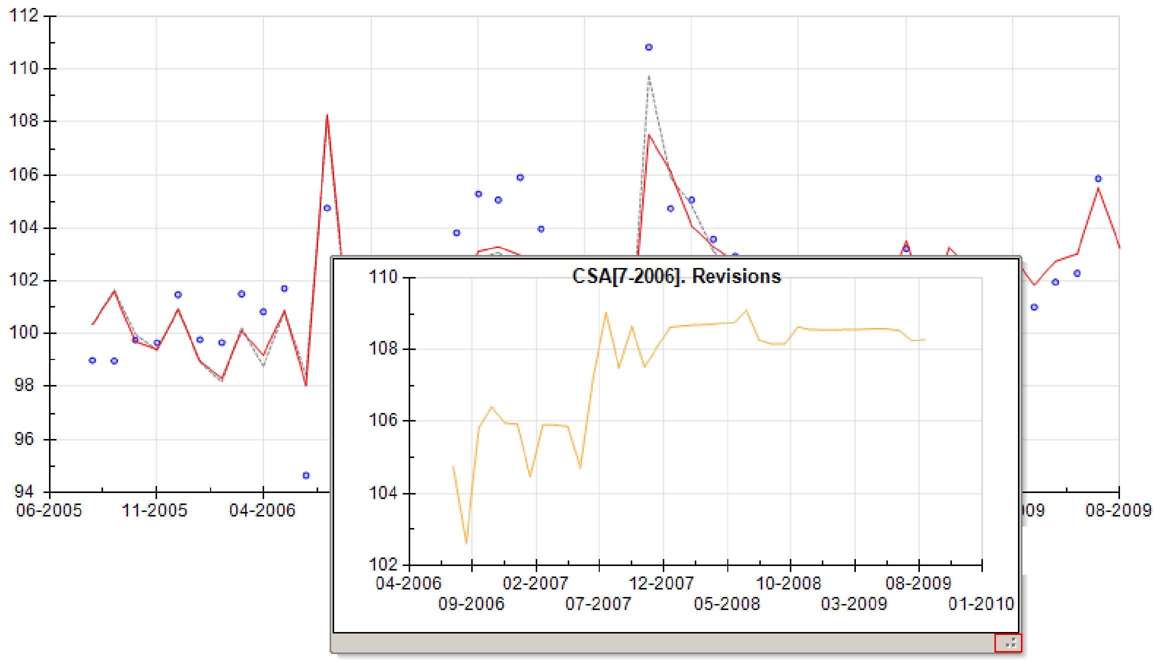
Revision history is stability diagnostic which visualise how a time series is affected when new observations are introduced. This statistics is generated both for SA series and trend-cycle component.

For each point the revision history shows the initial adjustment obtained when this point is the last observation in the time series (blue circle) and the later adjustment based on all available observations at present (red line). The difference between those two values is called a revision. As a rule, smaller revisions are better. Revision history is useful for comparing results from competing models. When the user defines two seasonal adjustment models for one time series and both these models are acceptable and then revision history can be used for choosing the better model in terms of revisions.

More detailed description is available in Annex.

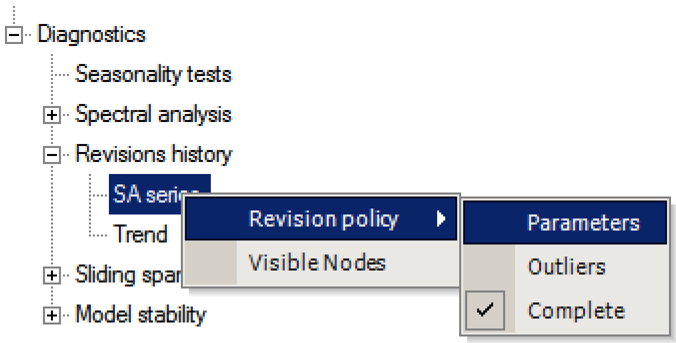



If the user clicks on a blue circle which represents the initial estimation for period $t_n$, an auxiliary window will appear. The figure shows the successive estimations (computed on $[t_0,...,t_n]$, $[t_0,...,t_{n+1}]$... $[t_0,...,t_T]$) of the considered series for the period $t_n$. From this figure the user can evaluate how the seasonally adjusted observations were changing from initial to final estimation. The analogous graph is available for trend analysis.

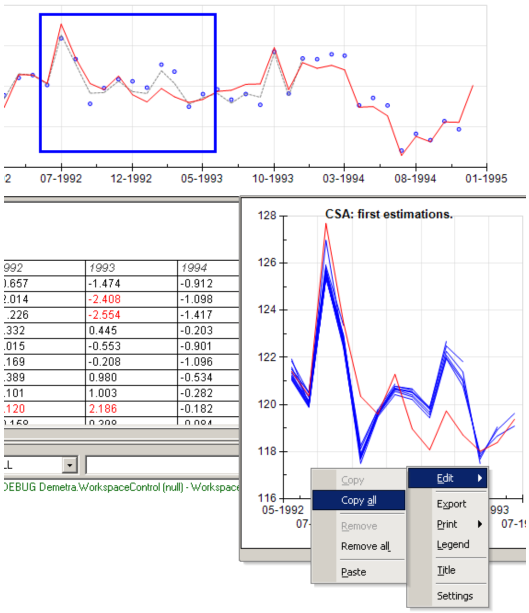



The user could establish the size of the revision using the vertical axis. In the figure above the revisions are about 5%. The figure size could be enlarged by dragging the right –bottom corner.

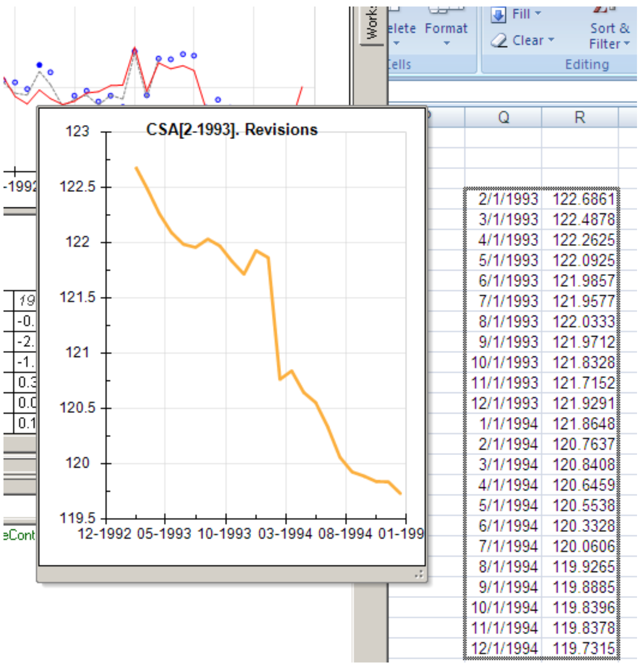
By default only the parameters of the model are re-estimated. It is also possible to make a complete re-estimation or a re-identification of the outliers. That option can be changed through the local menu of the revision history node (left panel), at the expense of the speed of the processing and for results that are usually very similar.




In the revisions history panels the user can have a complete overview of the different revisions for a given time span by selecting with the mouse (just like for zooming) the considered periods. The successive estimations are displayed in a separate pop-up window.




One can also get all the revisions for a specific period by clicking on the point that corresponds to the first estimate for that period. The results of those pop-up windows can be copied or dragged and dropped to other software (e.g. Excel).




The history analysis plot is accompanied by information about the relative difference between initial and final estimation for the last four years. Values which absolute value are larger (in absolute term) than 2 times the root mean squared error of the (absolute or relative) revisions are marked in red and provide information about the instability of the outcome.

For the additive decomposition absolute revisions are used, otherwise, relative differences are considered. The largest differences are displayed in red. They correspond to values that are larger (in absolute term) than 2 times the root mean squared error of the (absolute or relative) revisions.

**Relative differences**

*mean = 0,0679*
*rmse = 1,5881*

| | 2005 | 2006 | 2007 | 2008 | 2009 |
|---|---|---|---|---|---|
| January | | -0,818 | -1,986 | -0,272 | 0,460 |
| February | | -1,354 | -2,070 | -0,220 | 0,724 |
| March | | -1,383 | -1,680 | 0,516 | 0,567 |
| April | | -1,622 | -2,772 | 0,807 | 0,833 |
| May | | -0,810 | -2,465 | 0,660 | 0,775 |
| June | | 3,596 | -0,118 | -1,393 | 0,718 |
| July | | 3,366 | -1,159 | -1,190 | -0,320 |
| August | 1,367 | -0,122 | 3,615 | 1,244 | |
| September | 2,640 | 4,056 | 0,959 | 0,822 | |
| October | -0,060 | 0,346 | -2,984 | 0,282 | |
| November | -0,252 | -0,020 | 1,335 | 0,566 | |
| December | -0,559 | -0,946 | -0,932 | 0,509 | |

### *Sliding spans*

It is expected that seasonally adjusted data are stable, which means that removing or adding data points at either end of the series does not change the SA results very much. Sliding spans analysis is useful in case of seasonal brakes, large number of outliers and fast moving seasonality.

The sliding spans analysis checks the stability of SA. A span is a range of data between two dates. Sliding spans are series of two, three or four (depending on the length of the original time series, seasonal moving averages used and series' frequency) overlapping spans. The sliding spans analysis stands for the comparison of the correlated seasonal adjustments of a given observation obtained by applying the adjustment procedure to a sequence of three or four overlapping spans of data, all of which contain this observation. The procedure of withdrawing spans from time series is described in FINDLEY, D., MONSELL, B. C., SHULMAN, H. B., and PUGH, M. G. (1990). The program sets up 4 spans of 8 years, separated by 1 year. The seasonal and the trading day's panels compare the (relative) changes of the levels of those components. The SA changes panel is related to period changes. When an additive decomposition is used, the sliding spans analysis uses absolute differences. The threshold to detect abnormal values is set to 3% of the testing statistics.

The summary of Sliding spans analysis is presented below. It contains information about spans, results of the seasonality tests and means of seasonal factors for each month in each span. For the tests' description see Annex.

**Sliding spans summary**

**Time spans**

Span 1: from 1-1998 to 11-2006
Span 2: from 1-1999 to 11-2007
Span 3: from 1-2000 to 11-2008
Span 4: from 1-2001 to 11-2009

**Tests for seasonality**

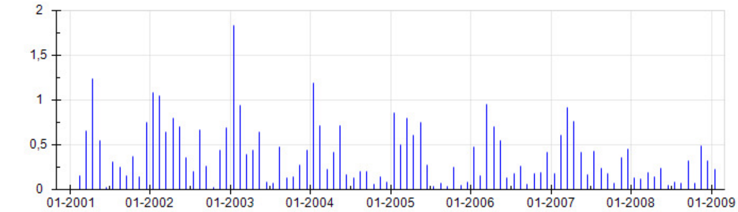
| | Span 1 | Span 2 | Span 3 | Span 4 |
|---|---|---|---|---|
| Stable seas. | 72,6 | 66,3 | 53,6 | 45,2 |
| Kruskal-Wallis | 68,6 | 74,4 | 77,5 | 76,3 |
| Moving seas. | 0,9 | 0,3 | 0,4 | 0,6 |
| Identifiable seas. | YES | YES | YES | YES |

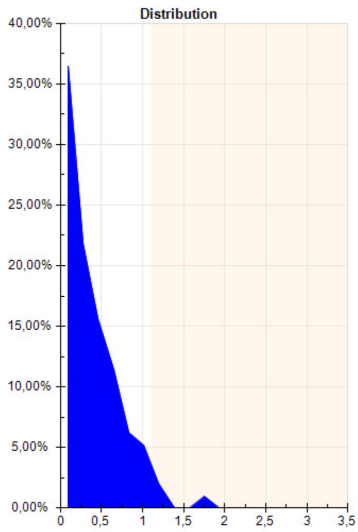
**Means of seasonal factors**

| | Span 1 | Span 2 | Span 3 | Span 4 |
|---|---|---|---|---|
| January | 1,0878 | 1,0697 | 1,0824 | 1,1212 |
| February | 0,8840 | 0,9067 | 0,9128 | 0,8949 |
| March | 0,8471 | 0,8513 | 0,8669 | 0,8489 |
| April | 0,8929 | 0,9149 | 0,9257 | 0,9322 |
| May | 0,7910 | 0,8006 | 0,7869 | 0,7692 |
| June | 0,7888 | 0,7623 | 0,7582 | 0,7752 |
| July | 0,9041 | 0,9240 | 0,9394 | 0,9650 |
| August | 0,8844 | 0,9050 | 0,9023 | 0,8854 |
| September | 0,8761 | 0,8776 | 0,8828 | 0,8726 |
| October | 1,0697 | 1,0984 | 1,1071 | 1,0856 |
| November | 1,1093 | 1,1037 | 1,0824 | 1,0920 |
| December | 1,8532 | 1,7882 | 1,7649 | 1,7665 |

Detailed results of sliding spans analysis conducted separately for seasonal component, trading days effect and SA series (changes), are presented in three graphs. Upper panel shows the sliding spans statistic for each period, the bottom-left panel presents the distribution of sliding spans unstable periods (months or quarters). Bottom-left panel contains detailed information about the percentage of values for which sliding spans condition is not fulfilled. It gives idea weather observations with unreliable adjustment cluster in certain calendar periods.

According to the FINDLEY, D., MONSELL, B. C., SHULMAN, H. B., and PUGH, M. G. (1990), the results' of seasonal adjustment are stable if the percentage of unstable (abnormal) seasonal factors is less than 15%. Empirical surveys support the view that adjustments with more than 25% of the months (or quarters) flagged for unstable seasonal factor estimates are not acceptable.




Abnormal values : 3,1%

**Breakdowns of unstable factors and Average Maximum Differences across spans**

| Period | Breakdowns | Average |
|---|---|---|
| **January** | **2** | **0,7** |
| February | 0 | 0,5 |
| March | 0 | 0,6 |
| **April** | **1** | **0,6** |
| May | 0 | 0,6 |
| June | 0 | 0,2 |
| July | 0 | 0,2 |
| August | 0 | 0,3 |
| September | 0 | 0,2 |
| October | 0 | 0,1 |
| November | 0 | 0,3 |
| December | 0 | 0,4 |

| Year | Breakdowns | Average |
|---|---|---|
| **2001** | **1** | **0,4** |
| 2002 | 0 | 0,6 |
| **2003** | **1** | **0,5** |
| **2004** | **1** | **0,4** |
| 2005 | 0 | 0,4 |
| 2006 | 0 | 0,4 |
| 2007 | 0 | 0,4 |
| 2008 | 0 | 0,2 |
| 2009 | 0 | 0,2 |

If the result of the Sliding spans analysis reveals many unstable estimates, it can support an idea of changing the model's specification. The example of such situation is presented below. Because of the large share of moving seasonality, the test for presence of identifiable seasonality failed.

**Sliding spans summary**

**Time spans**

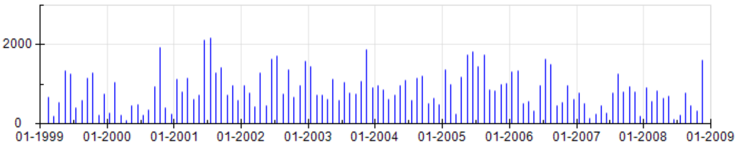
Span 1: from 1-1998 to 11-2006
Span 2: from 1-1999 to 11-2007
Span 3: from 1-2000 to 11-2008
Span 4: from 1-2001 to 11-2009

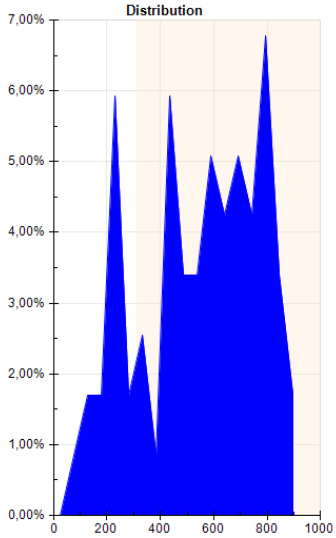
**Tests for seasonality**

| | Span 1 | Span 2 | Span 3 | Span 4 |
|---|---|---|---|---|
| Stable seas. | 6,7 | 6,6 | 7,0 | 7,3 |
| Kruskal-Wallis | 50,9 | 51,2 | 55,1 | 52,7 |
| Moving seas. | **4,4** | **4,7** | **4,8** | **4,1** |
| Identifiable seas. | **NO** | **NO** | **NO** | **NO** |

**Means of seasonal factors**

| | Span 1 | Span 2 | Span 3 | Span 4 |
|---|---|---|---|---|
| January | 221,813 | 97,6732 | 78,8923 | 143,397 |
| February | -1521,15 | -1508,5 | -1473,18 | -1376,74 |
| March | -1421,88 | -1632,14 | -1742,64 | -2038,82 |
| April | -966,52 | -700,687 | -418,783 | -510,623 |
| May | -1682,99 | -2182,04 | -2264,85 | -2265,83 |
| June | -960,968 | -949,774 | -1034,25 | -954,905 |
| July | 111,038 | 341,448 | 443,278 | 542,392 |
| August | -1055,2 | -1064,05 | -870,327 | -847,088 |
| September | -1135,11 | -864,588 | -713,117 | -1129,7 |
| October | 569,255 | 973,422 | 1102,41 | 1001,97 |
| November | 373,59 | 214,519 | 10,2962 | 421,368 |
| December | 7413,61 | 7305,25 | 6934 | 7034,75 |




Abnormal values : 88,1%

Breakdowns of unstable factors and Average Maximum Differences across spans

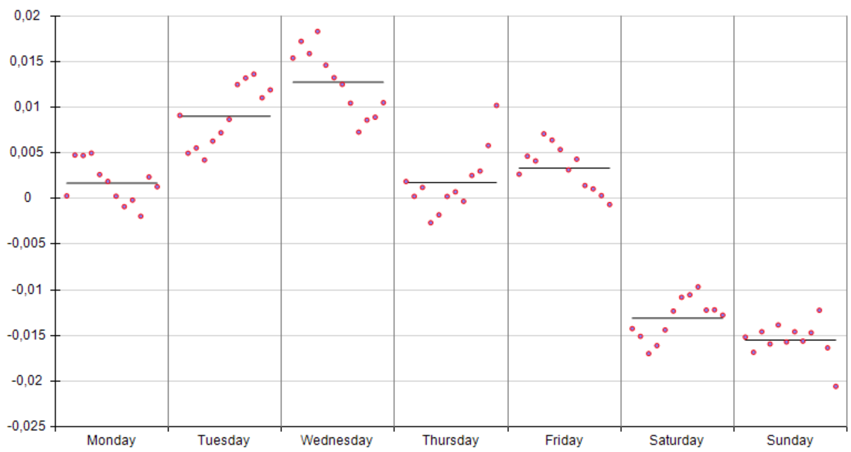
| Period | Breakdowns | Average |
|---|---|---|
| January | 8 | 1016,4 |
| February | 10 | 829,1 |
| March | 6 | 506,5 |
| April | 8 | 652,0 |
| May | 10 | 824,5 |
| June | 8 | 1035,5 |
| July | 8 | 1024,3 |
| August | 10 | 1017,7 |
| September | 10 | 937,3 |
| October | 10 | 880,0 |
| November | 9 | 945,5 |
| December | 7 | 708,8 |

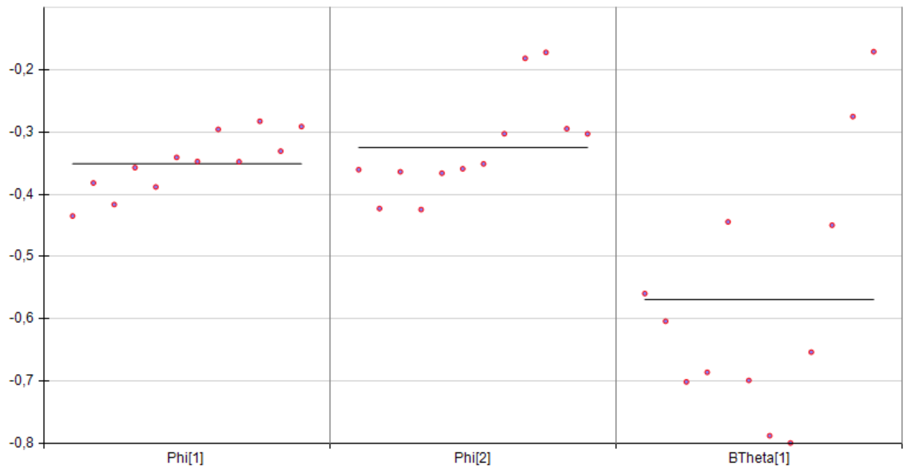
| Year | Breakdowns | Average |
|---|---|---|
| 1999 | 9 | 767,6 |
| 2000 | 7 | 555,0 |
| 2001 | 12 | 1142,5 |
| 2002 | 12 | 1047,8 |
| 2003 | 12 | 975,1 |
| 2004 | 12 | 813,8 |
| 2005 | 11 | 1186,0 |
| 2006 | 12 | 893,1 |
| 2007 | 8 | 596,5 |
| 2008 | 9 | 644,6 |

### *Model stability*

The diagnostics output window provides some purely descriptive features to analyze the stability of some part of the model, like trading days, Easter and Arima. *Model stability* analysis calculates ARIMA parameters and coefficients of the regressors for different periods and visualizes these results on the graphics. The parameters of the model chosen for the complete time span are computed on a moving window. The length of the window is 8 years. The points displayed on the figure correspond to the successive estimations. The figures are helpful for judging about the stability of the model parameters.

On the picture below the results of model stability diagnostic are shown.






#### 4.3.2.2 TramoSeats

The basic output structure is as follows:

- Main results,
  - Charts,
  - Table,
  - S-I ratio,
- Pre-processing (Tramo),
  - Pre-adjustment series,
  - Arima,
  - Regressors,
  - Residulas,
- Decomposition (Seats),
  - Stochastic series,
  - Model-based tests,
  - WK analysis,
- Diagnostics,
  - Seasonality tests,
  - Spectral analysis,

- Revisions history,
- Sliding spans,
- Model stability.

Detailed description of the seasonal adjustment outcomes is presented below. Because the majority of features are very similar to the X-12-ARIMA, the appropriate drawings are omitted. The user can find them in Seasonal adjustment results for X-12-ARIMA. In this section only those issues specific for TramoSeats will be discussed in details.

#### 4.3.2.2.1 Main results

Basic information about seasonal adjustment and the quality of the outcomes are presented in the following way.

trend. Innovation variance = 0,0966
seasonal. Innovation variance = 0,0372
irregular. Innovation variance = 0,2663

Diagnostics

**summary**
Good

**basic checks**
definition: Good (0,000)
annual totals: Good (0,004)

**visual spectral analysis**
spectral seas peaks: Bad
spectral td peaks: Good

**regarima residuals**
normality: Uncertain (0,036)
independence: Good (0,290)
spectral td peaks: Good (0,176)
spectral seas peaks: Bad (0,008)

**residual seasonality**
on sa: Good (1,000)
on sa (last 3 years): Good (0,983)
on irregular: Good (0,866)

**outliers**
number of outliers: Good (0,010)

**seats**
seas variance: Good (0,394)
irregular variance: Good (0,589)
seas/irr cross-correlation: Good (0,434)

Additional information is available in three subsections: *Charts*, *Table* and *S-I ratio.*

In *Charts* section the user will find:

- the original series with forecasts,
- the final seasonally adjusted series,
- the final trend with forecasts,
- the final seasonal component with forecasts,
- the final irregular component,
- the final seasonal component with forecasts.

The same time series are presented in *Table* section. The final estimation of the seasonal-irregular component and final seasonal factors are presented in the *S-I ratio* chart.

##### 4.3.2.2.2 Pre-processing

Table presented in pre-processing section contains series estimated by Tramo part. It includes interpolated series, series adjusted for calendar effects, deterministic component, calendar effects, trading days effect, outliers effect on irregular component and total outliers effect.

*Arima* section shows theoretical spectrum of the stationary and non-stationary model and autocorrelation function of the stationary part of the model.

*Regressions* section presents all regressors used in Tramo part, including trading days variables, leap year effect, outliers, Easter effect, ramps, intervention variables, user-defined variables.

In the next part the one-step ahead residuals from the model are presented both in the graph and the table. Analysis of the residuals consists of several tests and residuals' distribution. For details please refer to seasonal adjustment results for X-12-ARIMA and to Annex.

##### 4.3.2.2.3 Decomposition

The decomposition made by Seats assumes that all components in time series - trend, seasonal and irregular - are orthogonal and could be modeled using ARIMA model. Identification of the components requires that only irregular components include noise. ARIMA models estimated for each component are presented below:

**Decomposition**

Model
Non-stationary AR: 1 - B - B^12 + B^13
Stationary AR: 1
MA: 1 - 0,75112 B - 0,88252 B^12 + 0,66287 B^13
Innovation variance: 1,0000

trend
Non-stationary AR: 1 - 2 B + B^2
Stationary AR: 1
MA: 1 + 0,010354 B - 0,98965 B^2
Innovation variance: 0,0138

seasonal
Non-stationary AR: 1 + B + B^2 + B^3 + B^4 + B^5 + B^6 + B^7 + B^8 + B^9 + B^10 + B^11
Stationary AR: 1
MA: 1 + 0,67206 B + 0,38022 B^2 + 0,13433 B^3 - 0,060935 B^4 - 0,20492 B^5 - 0,30012 B^6 - 0,35125 B^7 - 0,36458 B^8 - 0,34727 B^9 - 0,30682 B^10 - 0,25071 B^11
Innovation variance: 0,0055

irregular
Non-stationary AR: 1
Stationary AR: 1
MA: 1
Innovation variance: 0,6779

Next section includes several tests. First of all variances of the component innovations are displayed (variance of the component innovation ("Component"), theoretical variances of the stationary transformation of the estimated components ("Estimator"), empirical variances of the stationary transformation of the estimated components ("Estimate"))[22]. SEATS identifies the components assuming that except from irregular they are clean of noise. It implies that the variance of irregular is maximized on the contrary the trend-cycle and seasonal component are stable as possible. The table compares the variance of the stationary transformation of the components (second column) with theirs estimators. The trend estimator always has a smaller variance and the ratio of the two variances get further away from one as the trend becomes more stable. Therefore, the more stochastic the trend is, the less will its variance be underestimated. On the other hand, the variations of a relatively stable will be extremely underestimated[23].

**Variance**

| | Component | Estimator | Estimate | PValue |
|---|---|---|---|---|
| trend | 0,4874 | 0,2290 | 0,3042 | 0,0242 |
| sa | 1,0718 | 0,5978 | 0,5553 | 0,5730 |
| seasonal | 5,1675 | 0,1904 | 0,1585 | 0,6777 |
| irregular | 0,0292 | 0,0049 | 0,0035 | 0,0446 |

For each component Demetra+ exhibits the values of the twelve consecutive lags (from lags-1 to lags-12) autocorrelations, its theoretical MMSE (minimum mean-squared error estimator) estimator and estimate actually obtained. Comparison of the theoretical MMSE estimator with the estimate actually calculated can be used as a diagnostic tool. The close agreement between estimator and estimate points towards validation of the results[24].

22 MARAVALL, A. (1995).
23 See MARAVALL, A. (1993).
24 GOMEZ, V., and MARAVALL, A. (2001).

**trend**

| Lag | Component | Estimator | Estimate | PValue |
|---|---|---|---|---|
| 1 | 0,0006 | 0,3491 | 0,3382 | 0,8163 |
| 2 | -0,4994 | -0,4096 | -0,3880 | 0,7499 |
| 3 | 0,0000 | -0,3233 | -0,2968 | 0,7332 |
| 4 | 0,0000 | -0,0805 | -0,0822 | 0,9869 |
| 5 | 0,0000 | -0,0198 | 0,0274 | 0,6648 |
| 6 | 0,0000 | -0,0040 | 0,0613 | 0,4618 |
| 7 | 0,0000 | 0,0024 | 0,0680 | 0,5454 |
| 8 | 0,0000 | 0,0139 | 0,0612 | 0,6498 |
| 9 | 0,0000 | 0,0564 | -0,0345 | 0,3619 |
| 10 | 0,0000 | 0,0715 | -0,1130 | 0,0710 |
| 11 | 0,0000 | -0,0607 | -0,1907 | 0,2158 |
| 12 | 0,0000 | -0,1740 | -0,1750 | 0,9919 |

**sa**

| Lag | Component | Estimator | Estimate | PValue |
|---|---|---|---|---|
| 1 | -0,5969 | -0,5986 | -0,6317 | 0,4856 |
| 2 | 0,0970 | 0,0988 | 0,1275 | 0,7643 |
| 3 | 0,0000 | 0,0000 | 0,0561 | 0,5775 |
| 4 | 0,0000 | 0,0000 | -0,1122 | 0,2650 |
| 5 | 0,0000 | 0,0000 | 0,0909 | 0,3468 |
| 6 | 0,0000 | 0,0000 | -0,0113 | 0,8972 |
| 7 | 0,0000 | 0,0000 | -0,0665 | 0,4909 |
| 8 | 0,0000 | 0,0000 | 0,1176 | 0,2430 |
| 9 | 0,0000 | 0,0000 | -0,0814 | 0,4194 |
| 10 | 0,0000 | -0,0172 | -0,0266 | 0,9260 |
| 11 | 0,0000 | 0,1043 | 0,1501 | 0,6312 |
| 12 | 0,0000 | -0,1742 | -0,2634 | 0,3057 |

**irregular**

| Lag | Component | Estimator | Estimate | PValue |
|---|---|---|---|---|
| 1 | 0,0000 | -0,3752 | -0,4328 | 0,3377 |
| 2 | 0,0000 | -0,0937 | -0,0444 | 0,5740 |
| 3 | 0,0000 | -0,0234 | 0,0319 | 0,5251 |
| 4 | 0,0000 | -0,0058 | -0,1027 | 0,2634 |
| 5 | 0,0000 | -0,0014 | 0,0812 | 0,3419 |
| 6 | 0,0000 | -0,0003 | 0,0010 | 0,9863 |
| 7 | 0,0000 | 0,0002 | -0,0462 | 0,5937 |
| 8 | 0,0000 | 0,0010 | 0,1110 | 0,2042 |
| 9 | 0,0000 | 0,0041 | -0,0777 | 0,3456 |
| 10 | 0,0000 | 0,0163 | -0,0266 | 0,6209 |
| 11 | 0,0000 | 0,0654 | 0,0845 | 0,8237 |
| 12 | 0,0000 | -0,1742 | -0,2383 | 0,3926 |

The decomposition made by SEATS assumes orthogonal components. The table below contains the correlations between the stationary transformations of the estimators and of the estimates

actually calculated by SEATS. The last column (PValue) displays the results of the test for no correlations between components. In the example below PValues are green, which indicates that all correlations are negligible.

Cross-correlation

| | Estimator | Estimate | PValue |
|---|---|---|---|
| trend/seasonal | -0,1121 | -0,0794 | 0,7135 |
| trend/irregular | -0,0936 | -0,0548 | 0,6909 |
| seasonal/irregular | 0,0446 | 0,0095 | 0,4344 |

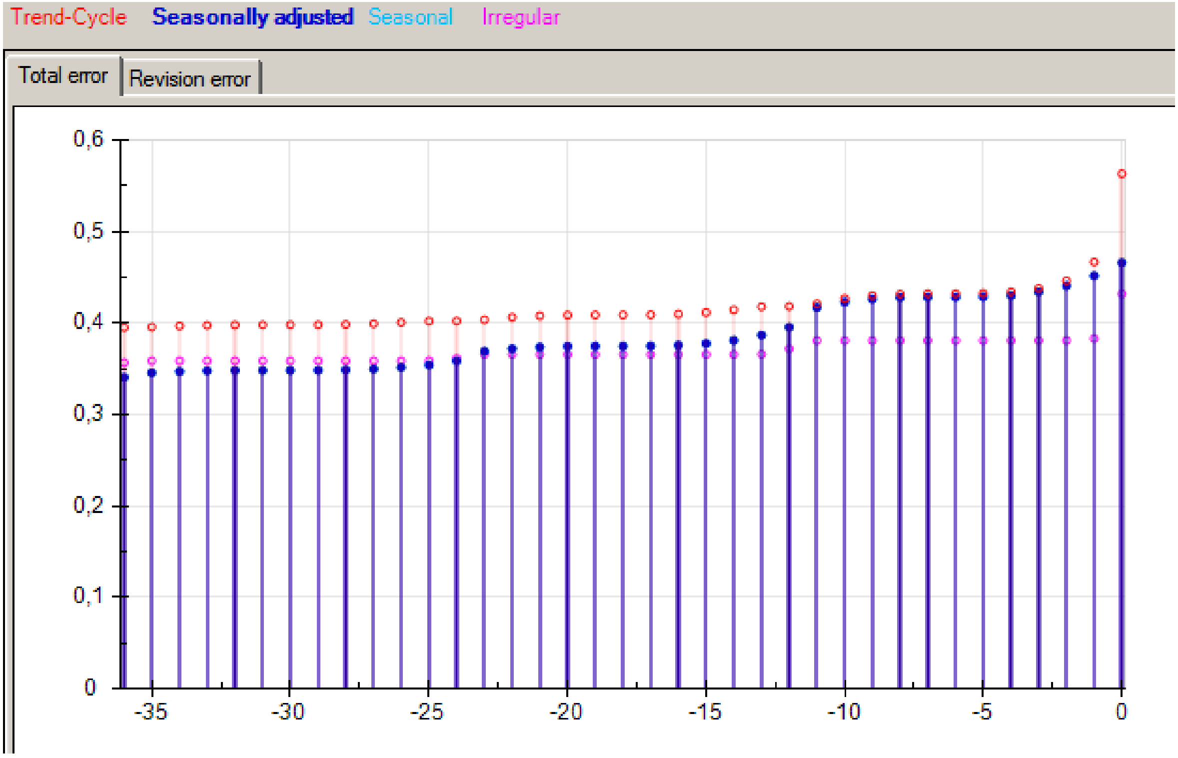
Additional information presented by Demetra+ is set of stochastic series (seasonally adjusted series, trend, seasonal component, irregular component, trend-forecast, seasonal component-forecast) and Wiener-Kolmogorow analysis.

Wiener-Kolmogorow analysis concentrates on[25]:

- Components (spectrum, ACGF),
- Final estimators (spectrum, square gain function, WK filters, ACFG, PsiE-weights),
- Preliminary estimators (Frequency response (square gain function, phase effect), WK filter, ACFG),
- Revision analysis (total error, revision error).

Revision analysis compares the variance of the different estimation errors for the historical estimators of the trend-cycle, seasonally adjusted series, seasonal and irregular. The graph shows the duration of the revision period, i.e. how many periods it takes for a new observation to no longer significantly affect the estimate.

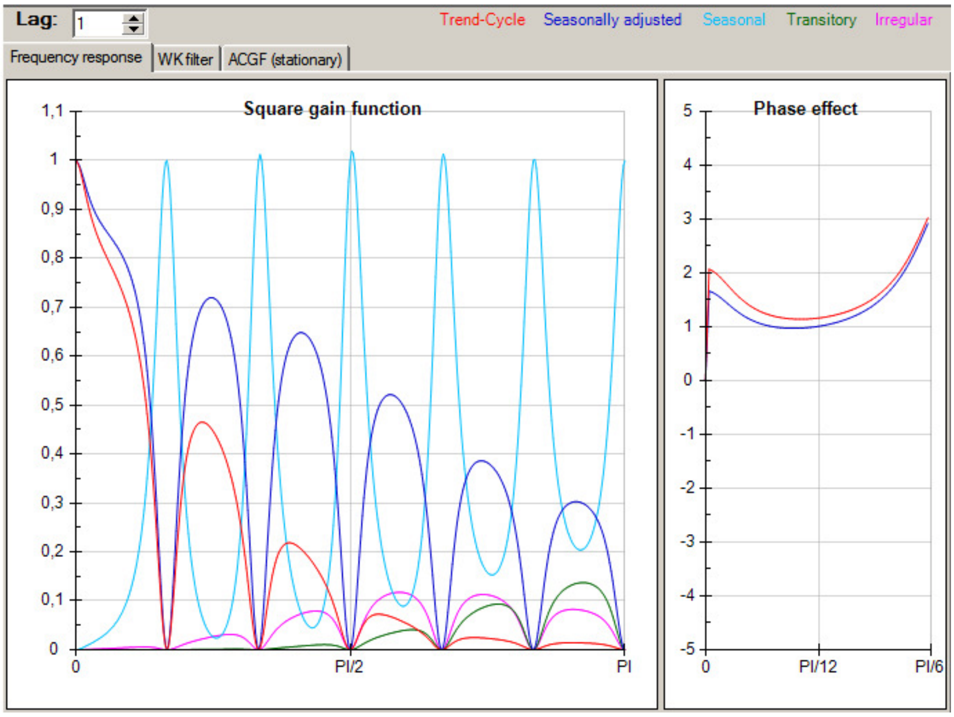
[25] Wiener-Kolmogorow analysis is described in: e.g. MARAVALL, A. (1993), MARAVALL, A. (2008), MARAVALL, A. (2006), MARAVALL, A. (1995).




Squared gains indicate which frequency components of the data are suppressed or amplified by the filter. Squared gain values larger than one suggest that the corresponding frequency component is stronger in the estimate than in the component, at least in the sense of contributing more variability.

Phase delays indicate how frequency components are shifted in time by the filter. Phase function is calculated for trend-cycle and seasonally adjusted series to evaluate how much seasonal filters delay business cycle information[26].

[26] See FINDLEY, D. F., MARTIN, D. E. K. (2006).




The “Decomposition” panel contains the ARIMA models which are defined by SEATS. The sub-panels of that part of the output present, for SEATS, many properties of the Wiener-Kolmogorov filters generated by the canonical decomposition.

#### 4.3.2.2.4 Diagnostics

Demerta+ offers the following seasonality tests:

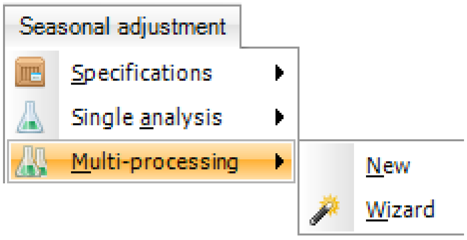
- Friedman test,
- Kruskal-Wallis test,
- Test for the presence of seasonality assuming stability,
- Evaluative seasonal test,
- Residual seasonality test,
- Combined seasonality test.

In the Diagnostic section the user will find also Spectral analysis, Revisions history, Sliding spans, Model stability. For details please refer to Seasonal adjustment results for X-12-ARIMA and to the Annex.

## 4.4 Multi-processing

Multi-processing specification is designed for quick and efficient seasonal adjustment of large data sets. Multi-processing specifications that mix different seasonal adjustment methods are available. The software provides two different ways to perform multi-processing. The first solution is based on the "active" specification; in that solution, the series that are subject to in a

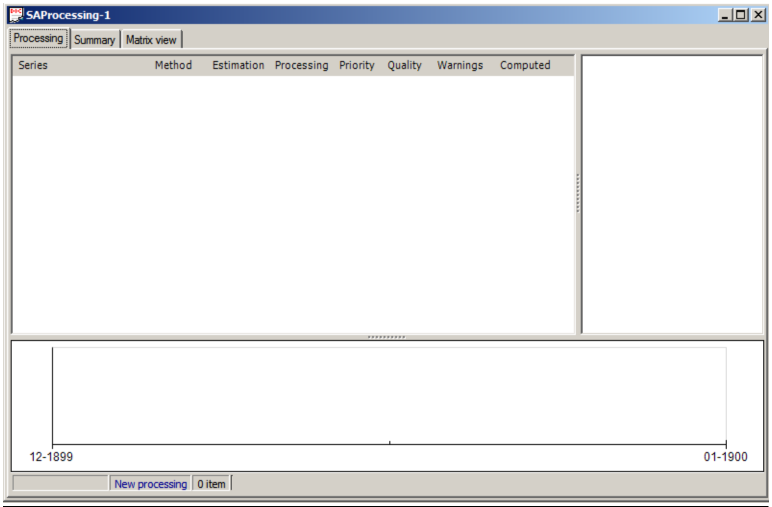
multi-processing are automatically associated with the "active" specification. The second solution consists in using a wizard, which allows the users to associate series and specifications step by step. Both functions are activated from the main menu.



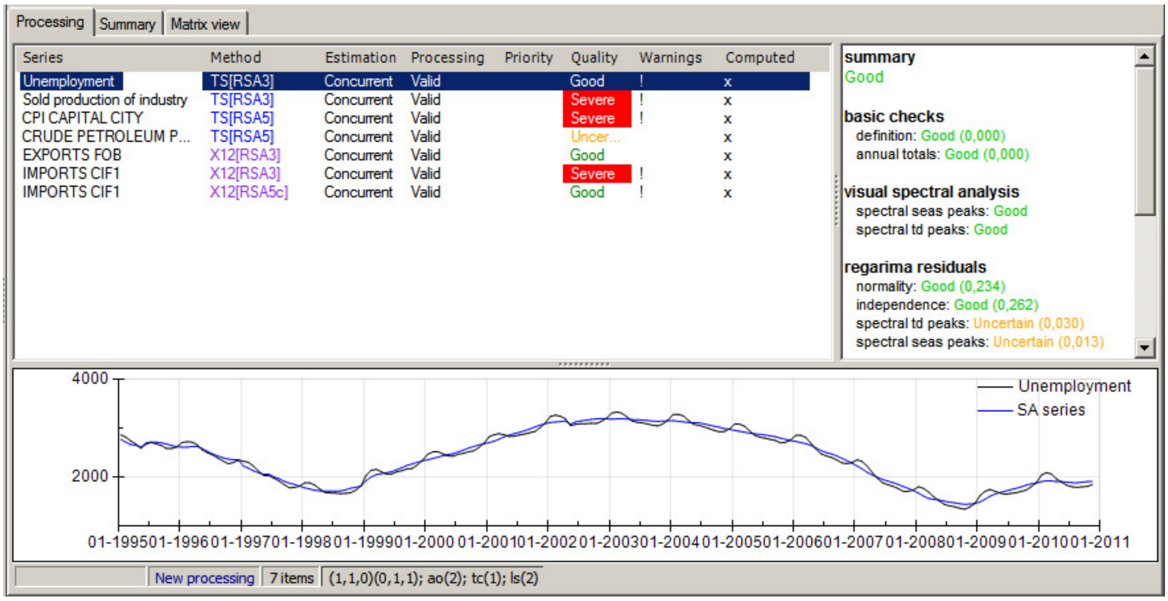

### 4.4.1 Defining a multi-processing

Creation of a new multi-processing

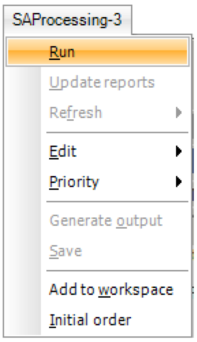
This option opens the following window.




The user should first activate the specification and then drag and drop the time series into the window. We recall that the active specification can be selected in the workspace through a local menu; it can be either a pre-defined specification or a user-defined one. The user can change his/hers choice of the active specification. It enables to launch the seasonal adjustment for one time series using different specifications in order to compare the results.




The processing is actually launched by means of the *Run* command under the *SAProcessing-1* main menu item.

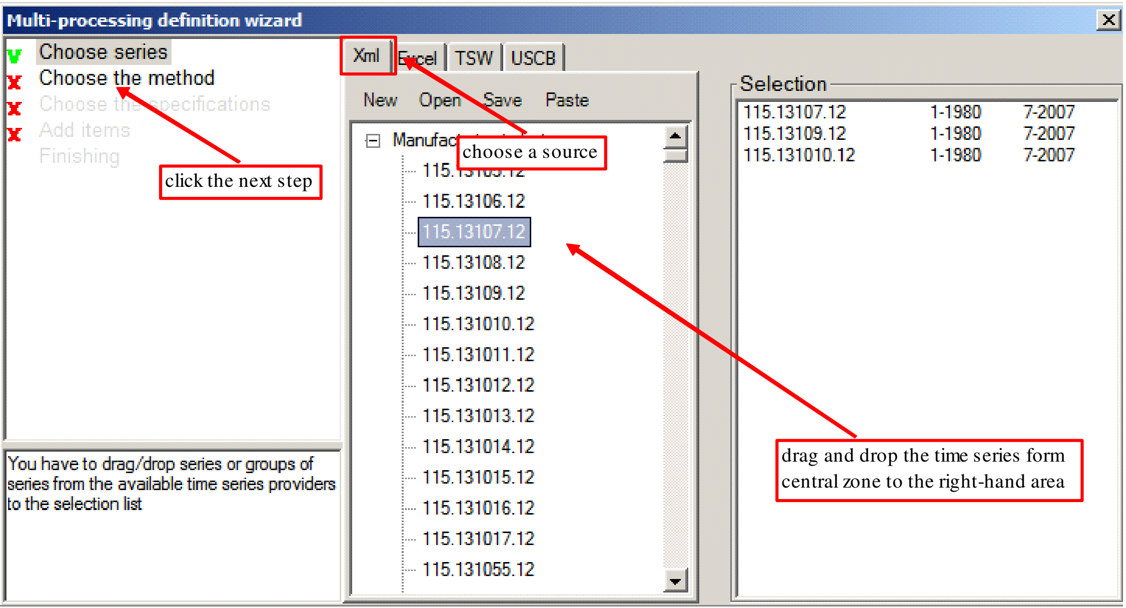



The user can also launch the seasonal adjustment of the time series by clicking on its name on the list.

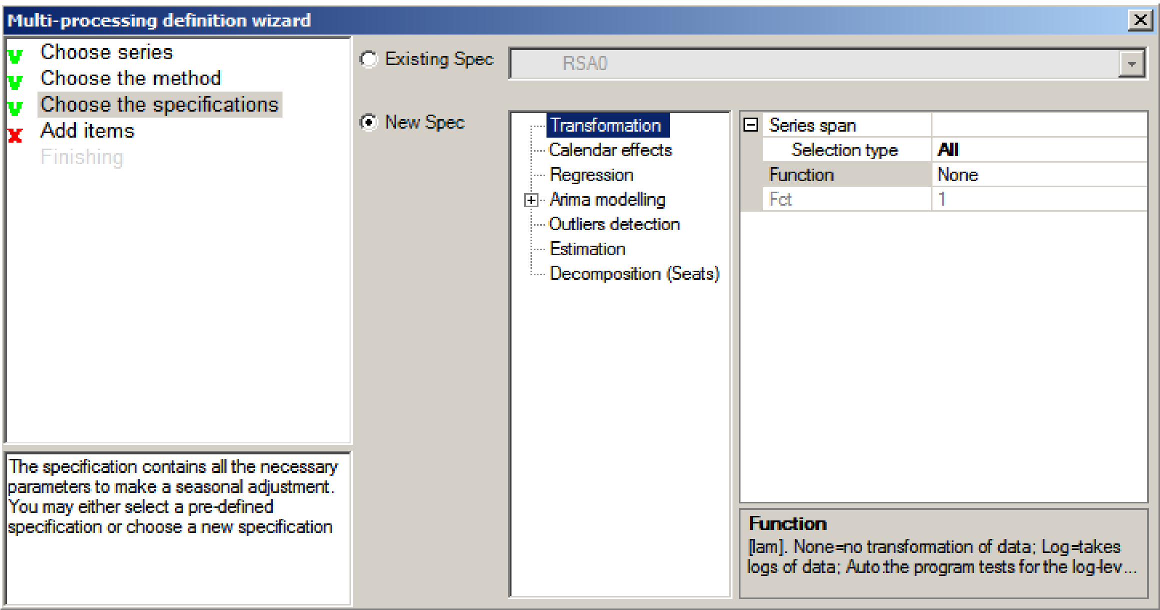
<u>Creation of a multi-processing via wizard</u>

When the user activates the wizard, the empty window is displayed. The wizard guides the user through the construction of the associations "series-specifications". It also gives him the possibility to define and to use specifications that don't belong to the workspace.
Consecutive steps are similar to those which were described in single seasonal adjustment part. However, there are two main differences.

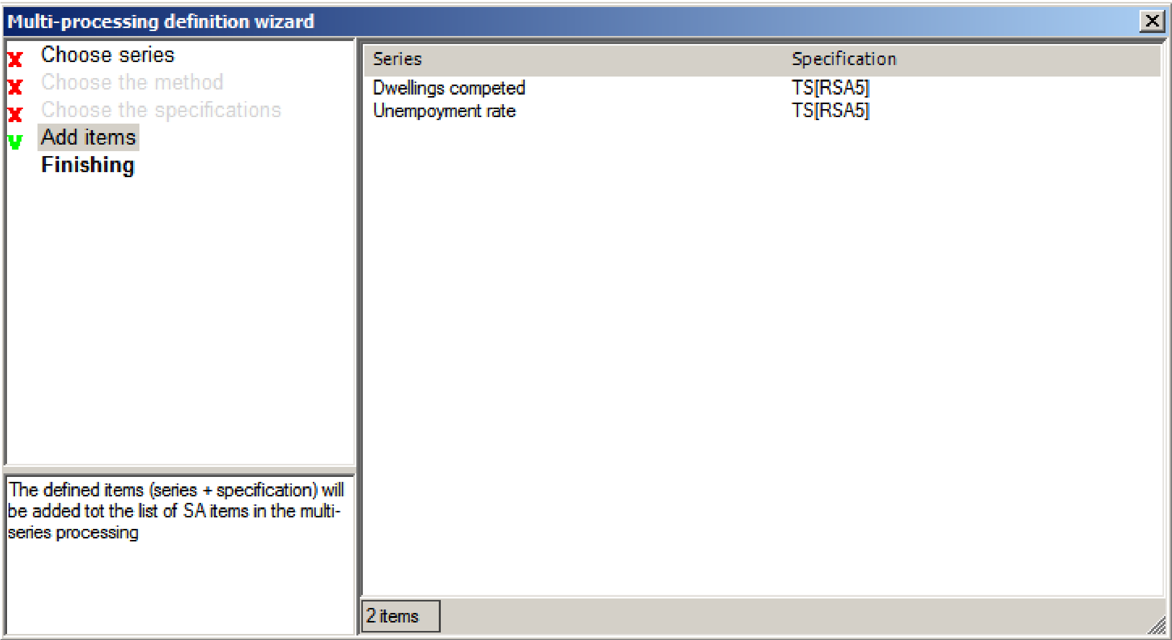
First of all, in the first panel the user can choose more than one time series and drop them into Selection window.




Then, the user should decide which seasonal adjustment method – X12 or TramoSeats - will be used. After that the user can chose existing specification or create new specification as it was shown in 4.1 and 4.2.




Next, in the add items Demetra+ presents time series which will be added to the list of seasonally adjusted items in the multi-series processing. Add items -part is not about adding time series to the regression part of the pre-adjustment model but simply shows the user the list of time series which have been chosen in the first step. It is not possible to add here new time series to the multi-processing.




At the last stage of the wizard ("Finishing") the user can modify the name of the multiprocessing (SAProcessing-xx, default); he can also add the multi-processing to his workspace, for future re-use and he can decide if the execution is automatically started (the default) when the wizard is closed. It should be mentioned that he can go back to the first step of the wizard at any time, if he wants to add other series with other specifications.

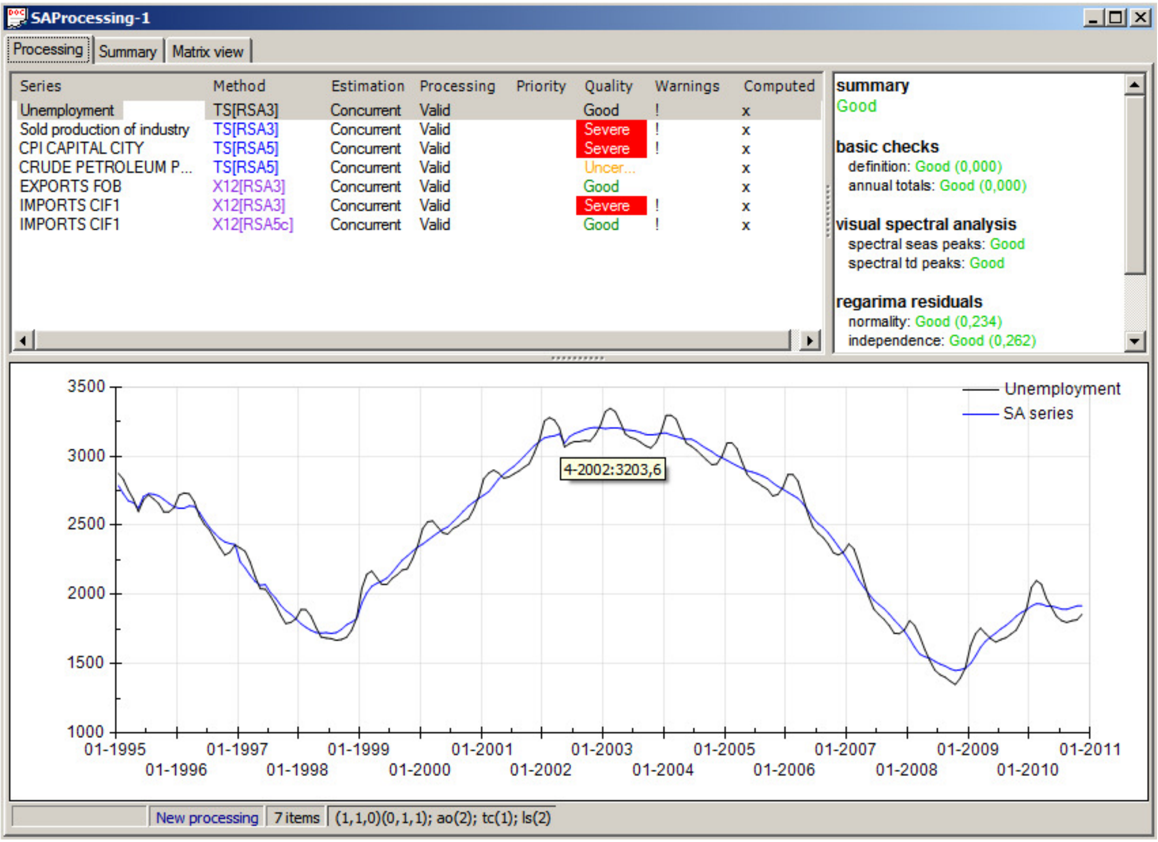
### 4.4.2 Seasonal adjustment results for multi-processing

#### *4.4.2.1 Generalities*

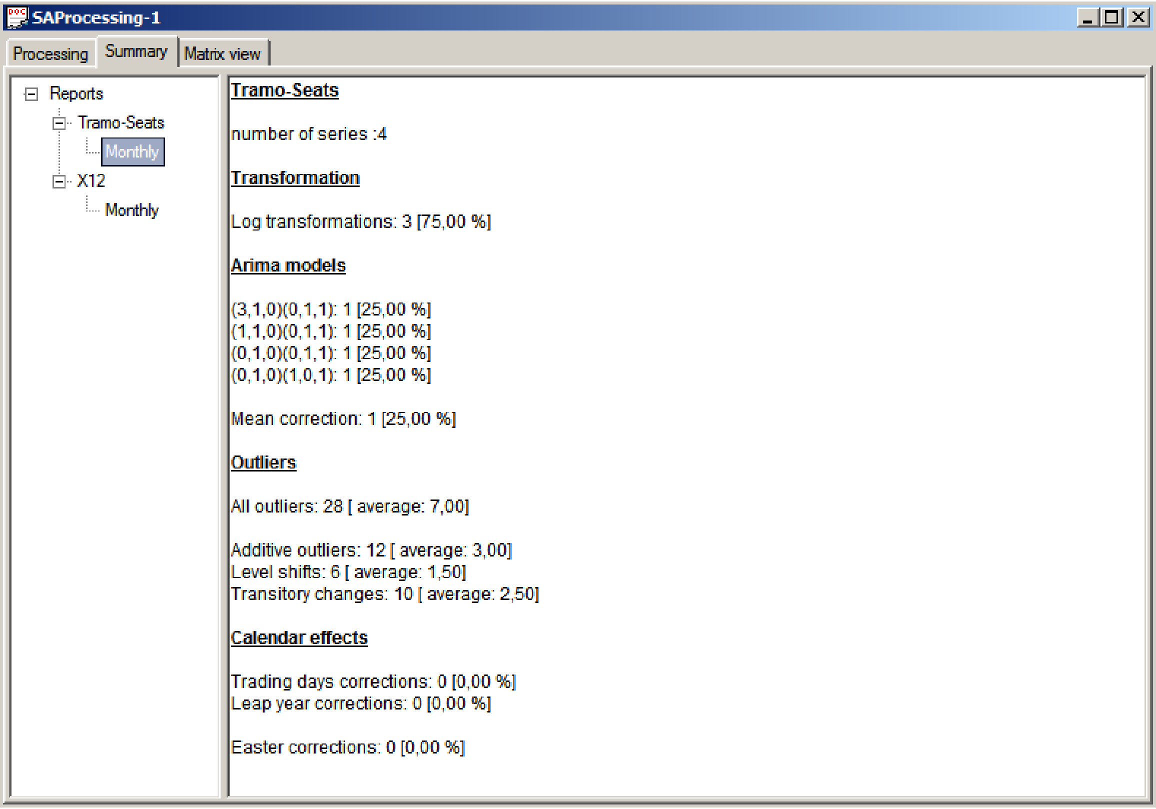
The outcome of the multi-processing is presented in the window which contains three panels.

The first panel - Processing - gives an overview of the processing of each series and more, especially of the diagnostics computed by Demetra+ on its seasonal adjustment. Some warnings can also be put forward, for short series, non-decomposable models (SEATS) or when the differenced series doesn't show seasonal peaks. Information on those warnings is displayed by a tooltip on the series. The user can sort the multi-processing by clicking a column header.

By clicking on the time series' name a summary of the tests' results is displayed in the right panel. For the description of those tests please refer to Chapter 4.3.2.1. At the bottom of the window the graph of final seasonally adjusted series and raw series is displayed.



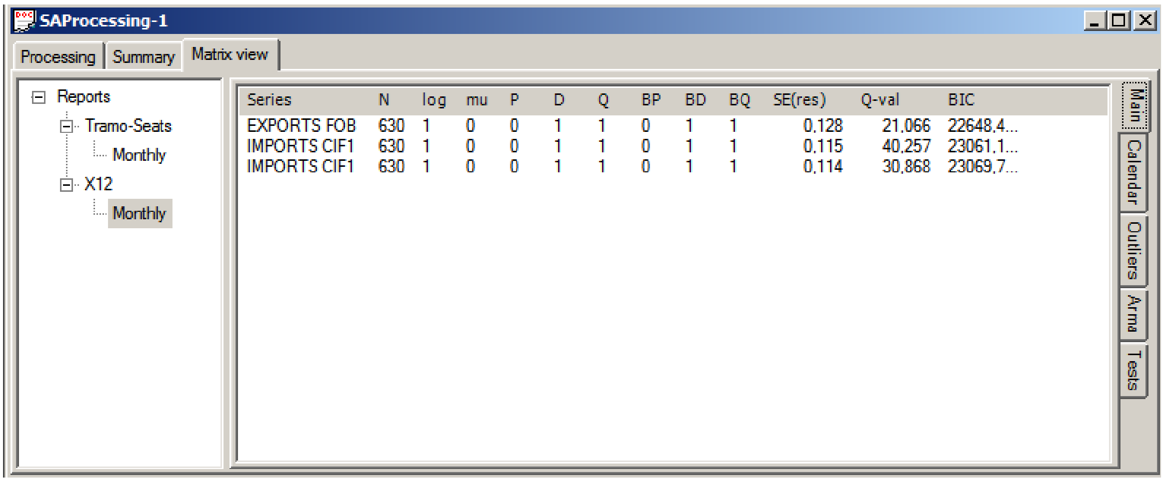

The *Summary* panel gives general information on the results obtained from each method for each frequency. The example below shows that TramoSeats method has been chosen for four time series. Three of them have been logarithmically transformed. The list of the Arima models shows the model parameters used in time series set. There were 28 outliers detected, the majority of which were additive outliers. Calendar effects weren't detected for any of the time series seasonally adjusted using TramoSeats method.




Last section – *Matrix view* panel – provides information similar to the matrix output of TSW (TramoSeats for Windows).

The summary information is divided into five folds available in the right side of the panel:

- **Main** – contains main statistical properties of the ARIMA model used in Pre-processing;
- **Calendar** – presents calendar specification results,
- **Outliers** – outlier structure of each series and coefficients of ARIMA model and their significance levels,
- **Arima** – parameters' values and theirs t-stat values,
- **Tests** – p-values of different tests computed on the residuals and with other information (annual discrepancies between raw and adjusted data, spectral visual peaks).




The matrices can be copied into the clipboard by the usual keys combination (Ctrl+C), for user in other software, like Excel.

#### 4.4.2.2 Multi-processing menu

Menu offers the following options for multi-processing:

**Run** – runs the defined multi-processing seasonal adjustment,

**Update reports** – updates the processing after changes in seasonal adjustment specifications,

**Refresh** – refreshing a processing with new data,

**Edit** – allows adding new times series to the list (using multi-processing wizard) and pasting previously cut time series again in the list. Last three edit options: cut, copy and delete are active if the time series was marked on the list,

**Priority** – offers two options: level-based and log-based,

**Save** – saves the processing,

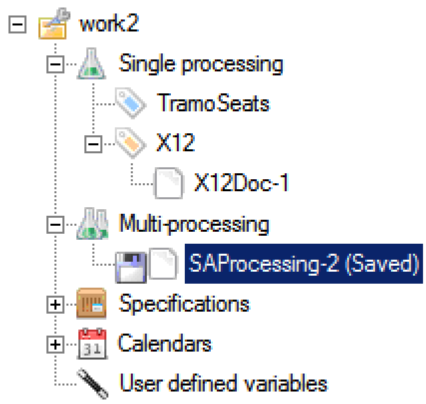
**Generate output** – offers many output formats (txt, XLS, ODBC, CSV, CSV matrix), the choice of the folder that will contain the results (in the example below the file will be saved on disk C:\Documents and Settings…) and the content of the exported file,

**Add to workspace** – adds the multi-processing to the workspace's tree,

**Initial order** – displays times series on the list in initial order. The option is useful if the list has been sorted by other column (e.g. by quality or method).

After defining a multi-processing the user should **run** the estimation. After that it is possible to **generate output**. The **save** option is inactive as soon as the user **add**s the processing **to** the

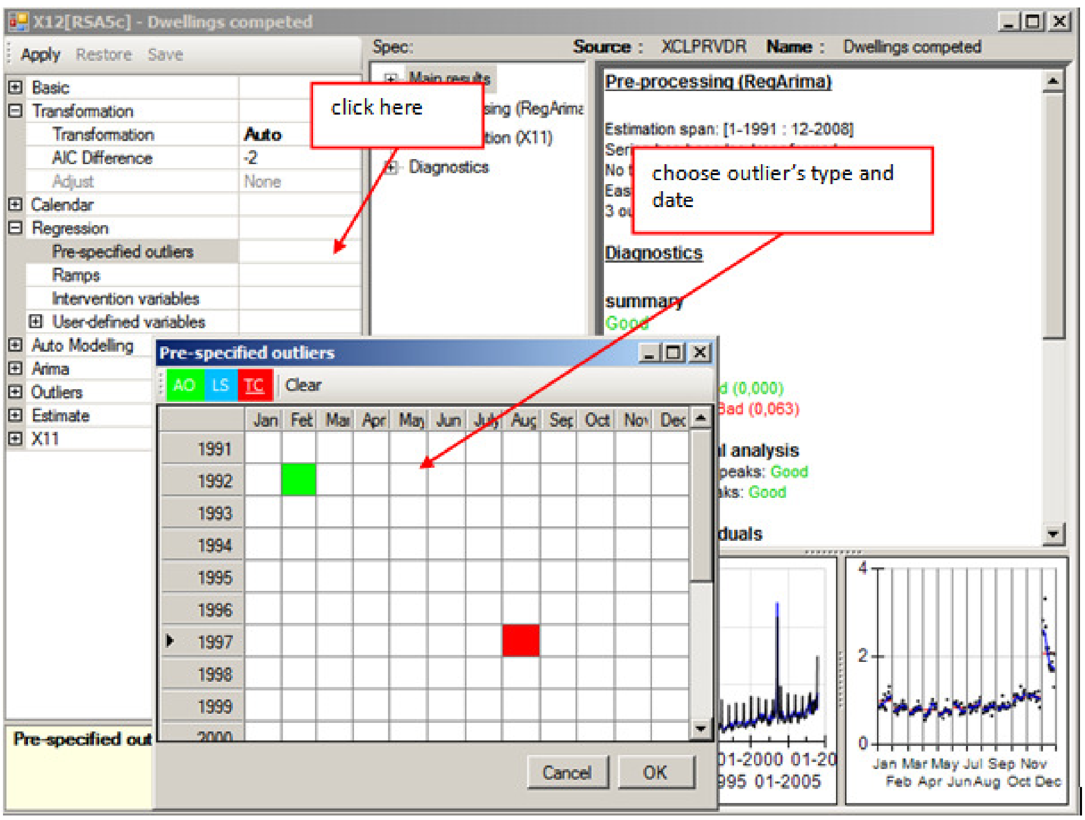
**workspace**. Once the output was created, the user can save the multiprocessing. The appropriate item will appear in the workspace tree.



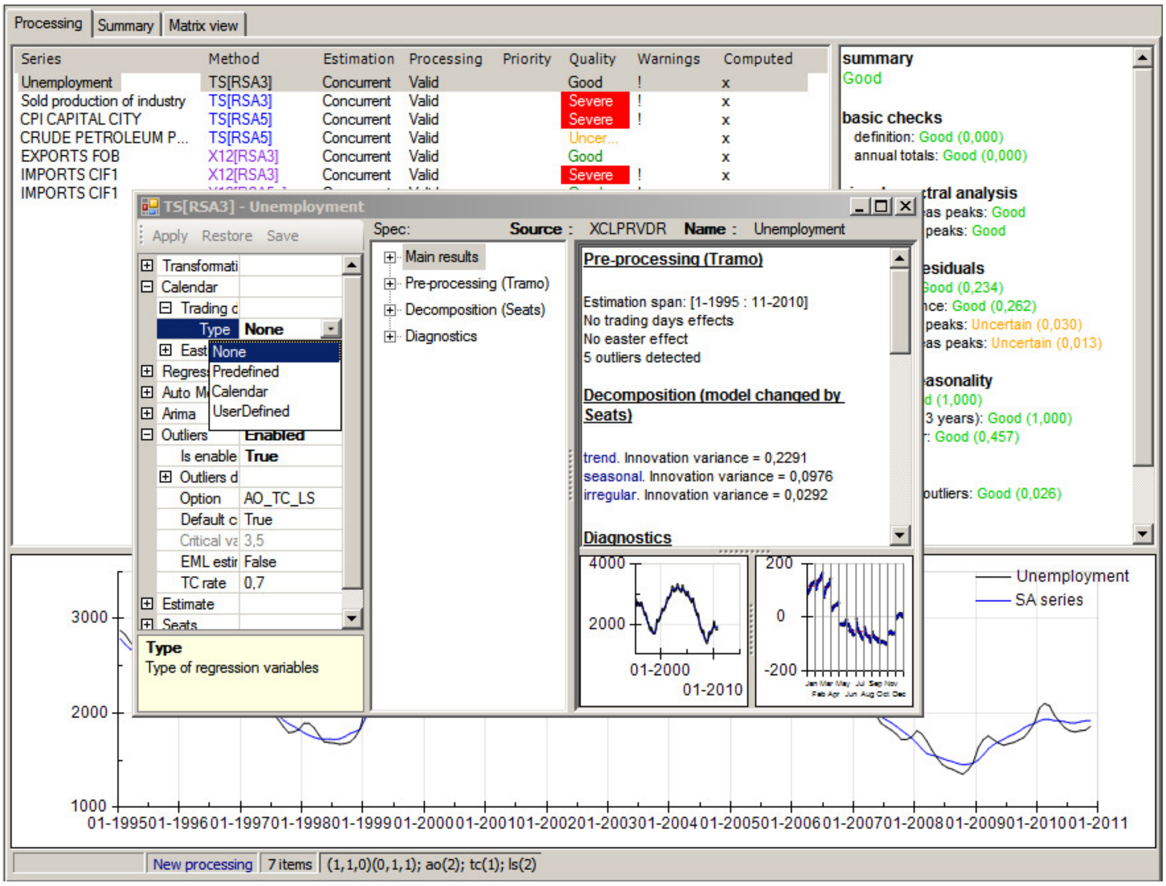

### 4.4.2.3 Detailed results

For each time series from multi-processing seasonal adjustment Demetra+ offers the access to the complete description of the seasonal adjustment results by a double click on the time series' name. This option is available for both *Processing* and *Matrix view* panels. The user can modify then the specification by changing the options in the left part of the window. This option could be useful in case the quality of a specific processing is insufficient and the user wishes to modify some options to get a better result.

As an example, the following panel shows how to change the pre-specified outliers.



When the new options are chosen, the user should click on **Apply** button to launch the seasonal adjustment with modified settings.

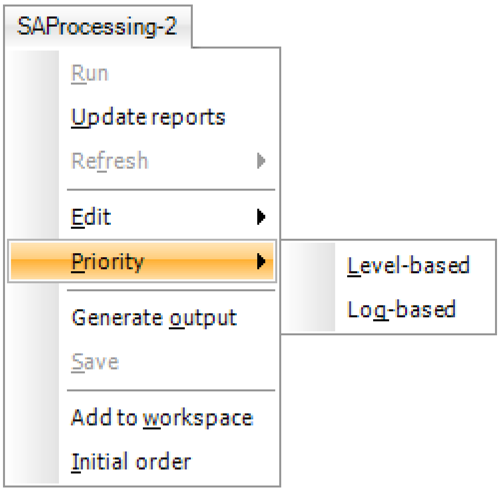


If the result is acceptable, the user can save it to the multi-processing window using **Save** button. The multi-processing contains now the adjusted specification for that series. Otherwise, the user can come back to the previous settings using **Restore** button.

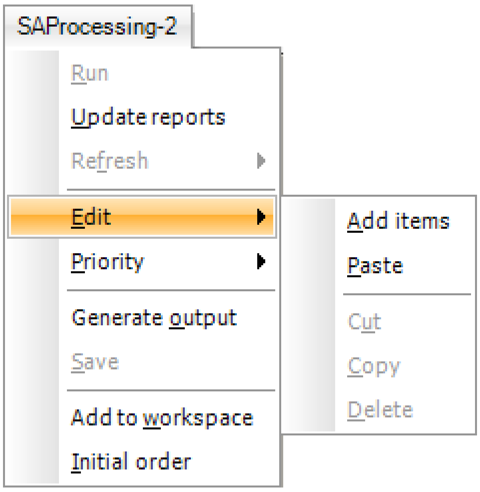
It is not necessary to close the details window to get information on another series; that window is updated by a simple click on another series of the multi-processing view.

It is also possible to create a separate single-processing from a multi-processing document by dragging the corresponding item from the series column to the central panel of Demetra+.

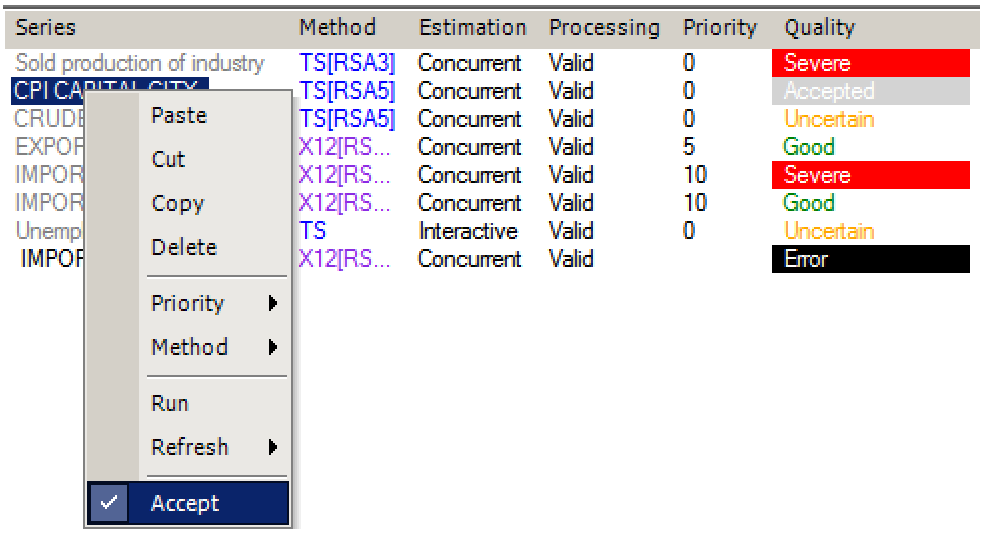
By means of right-click menu the user can paste, cut, copy and delete time series marked. Also priority can be changed into log-based or level based.




The user can add new time series to the  multi-processing, using *Edit->Add items option*.

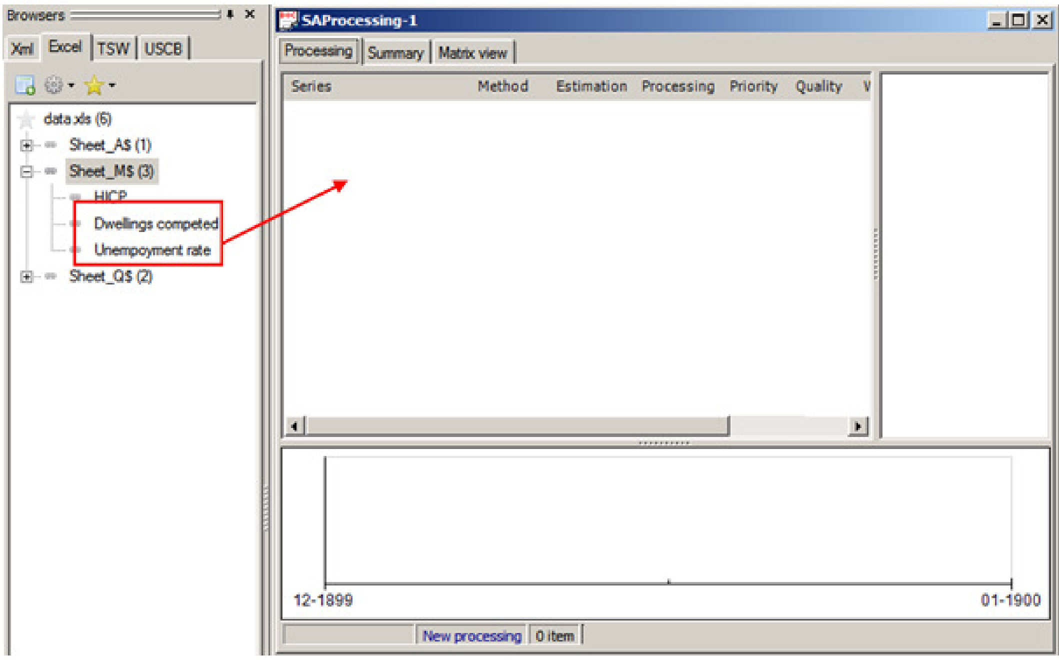



Demetra+ allows the user to accept the models, the quality of which wasn't satisfactory. If the user clicks on the *Accept* option, Demetra+ changes the message displayed in *Quality* column into *Accepted*.



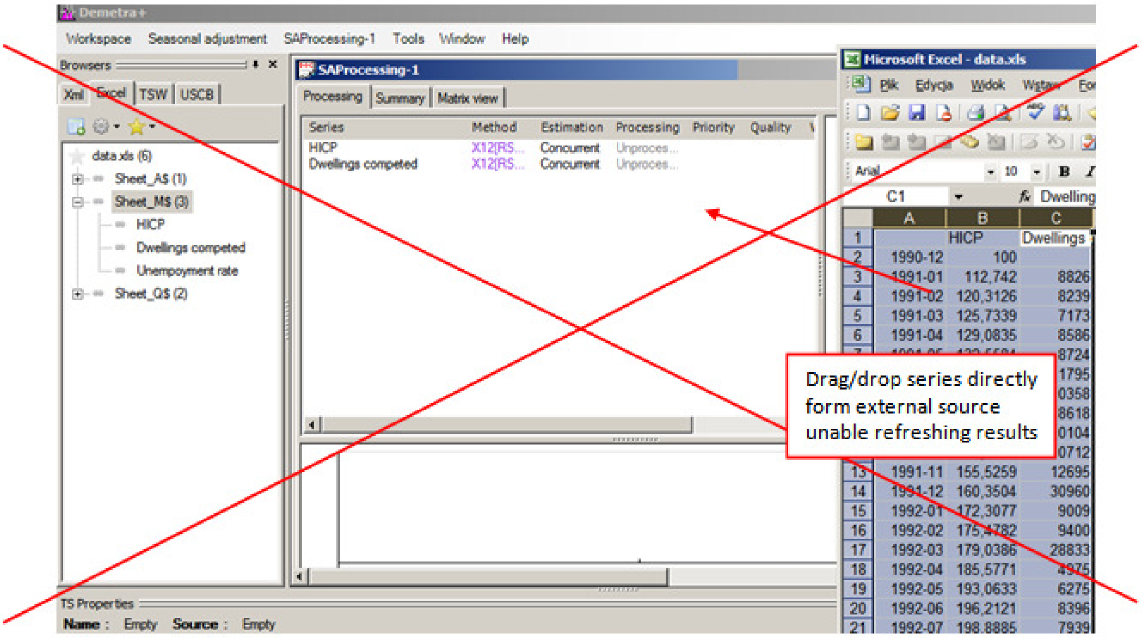

## 4.5 Period-to-period data production

Multi-processing is designed for regular production of the seasonally adjusted data. For this purpose the user should define multi-processing using the data from the browsers, i.e.:




In this case Demetra+ saves the location of the file from which the data come from.

If the variables in multi-processing come directly from external source, it won't be possible to update the processing. Such variables are static, so their location is not saved by Demetra+.

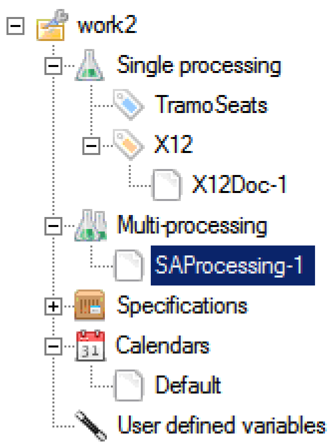



When the multi-processing was created, the user should added it to the workspace and then saved it using the options from multi-processing menu. Then the user can use this multi-processing for month-to-month (quarter-to-quarter) seasonal adjustment. This process should be conducted in the following way:

1. Update the time series in the external file or source from which the variables come from (e.g. update the file ‘data.xls’ with the new observations but don’t change neither the file’s name nor its location).

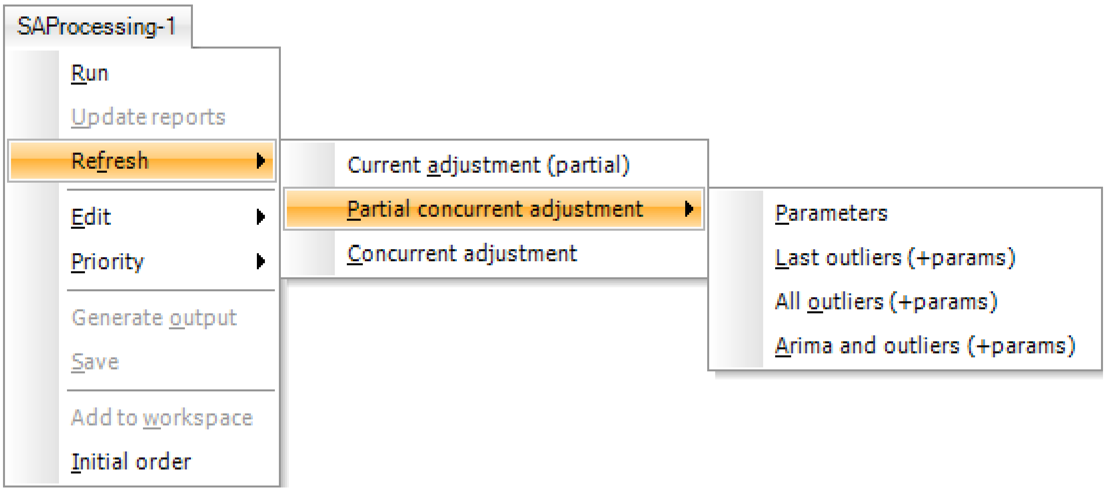
2. Start Demetra+.

3. Chose the multi-processing from the workspace tree by double-clicking on it.

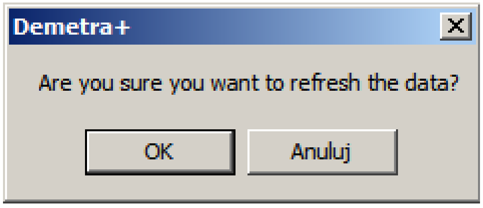



4. Choose in which way you would like to refresh the results[27].

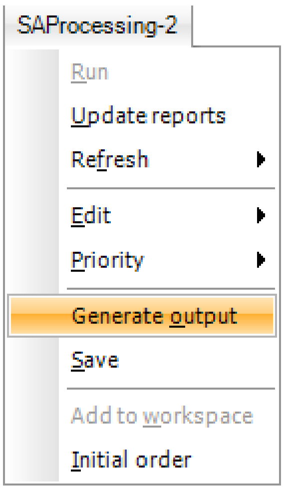
[27] For more details see 5.2.



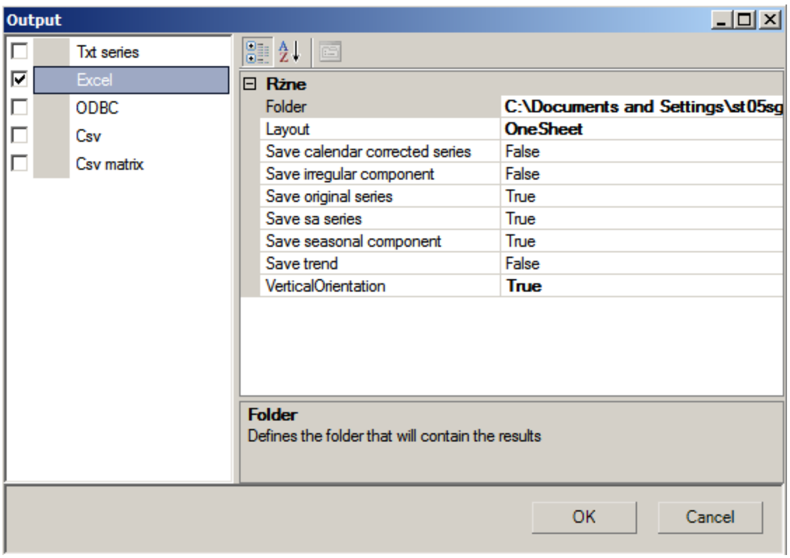

5. Confirm that you want to refresh the data.




6. Choose the option **Generate output** form the menu.




7. Mark the output and click **OK**.



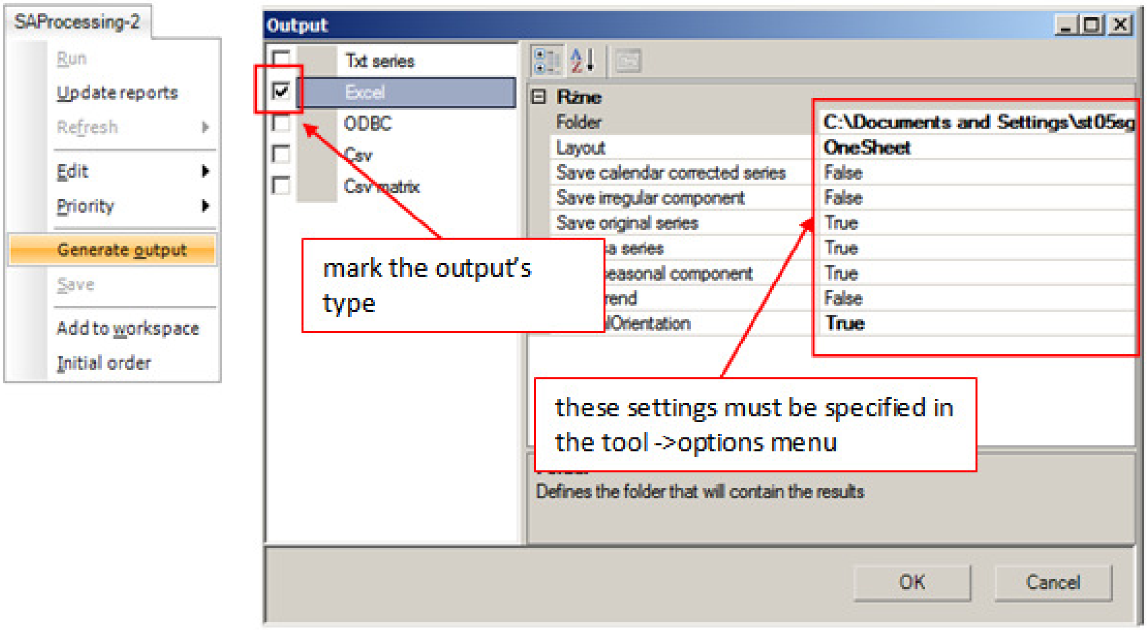

8. Demetra+ creates the file with the output. The old version of the file (e.g. filed created in the previous period) will be replaced by the new version.

Detailed aspects of saving the results in external files are discussed in the next section.

## 4.6 Sending the results to external devices

When the multi-processing is created, it is possible to generate several outputs (Excel workbook, csv files...), through the main menu command: *SAProcessingXXX -> Generate output* or (*TSProcessingXXX -> Generate output*). It should be noted that Excel and .csv outputs will be put in the temporary folder if their target folders are not specified.

The user is expected to choose the output format by marking the appropriate box in left-hand part of the *Output* window. The settings which are displayed in the second part of the window come from *Tool -> Options* menu. All changes in those settings should be done in the *Tool -> Options* menu. If the user changes the settings (e. g. output's folder) in the *SAProcessingXXX -> Generate output* window (or *TSProcessingXXX -> Generate output*), it will not have any effect on the output's content.




For multi-processing that don't belong to a workspace, output files' name is default ("demetra"). If multi-processing is saved in the workspace the multi-processing's name is used.

# 5 Additional functions

## 5.1 Changing the specification

The user is able to modify the used specification and to see immediately the result of changes made.

The specification is edited through the main menu: TramoSeatsDocxxx / X12Docxxx -> Specification... It is possible to edit the specification used to generate the processing (current specification) or the specification that corresponds to the results (result specification).

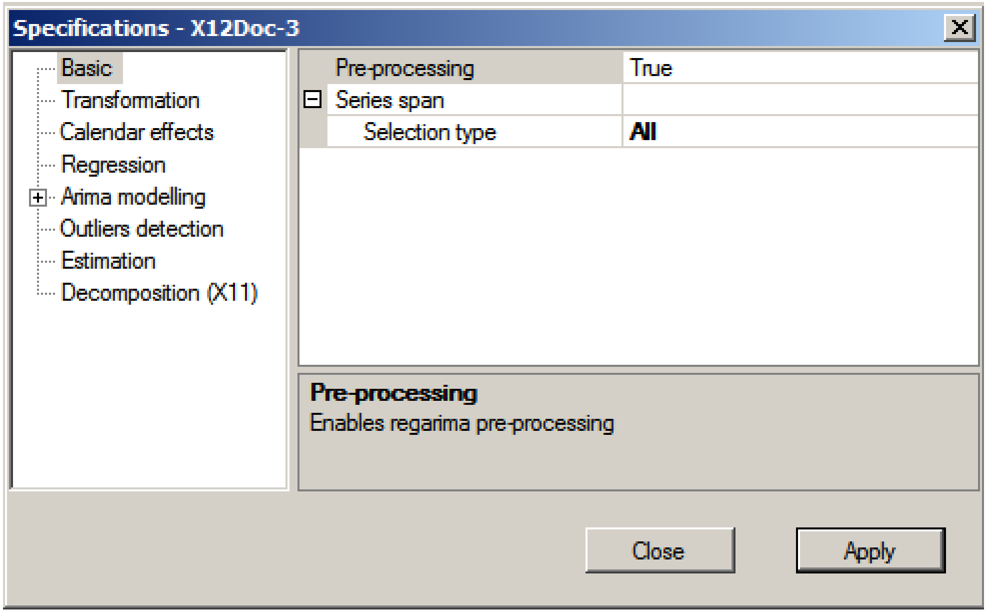
Current specification is displayed in a non modal dialog box, so the user can change any option and inspect its impact on the results. For a detailed description of the specifications, the user should refer to the 4.1 (X12) or 4.2 (TramoSeats).

The example below refers to X12:

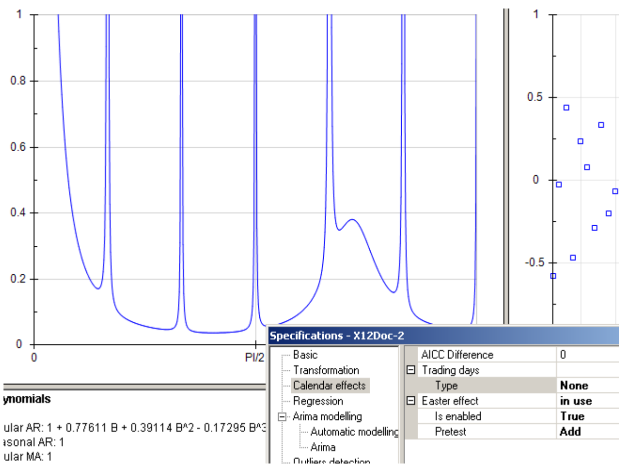
- Activate previously generated output from X12,
- Select from menu X12DocDocxxx-> Specification ->Current Specification,
- Modify the span of the series in the "Basic" panel:
  - Click on the Basic item in the left panel of the specification dialog box,
  - Expand the "series span" node in the right panel,

- Choose the "excluding" selection type,
- Write "12" in the "last" node,
- Press the "Apply" button.

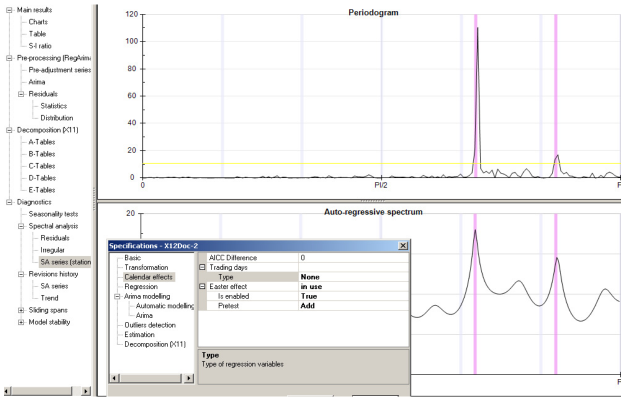
The processing is computed on the series without the last 12 observations. A visual comparison of the forecasts of X12 and of the actual figures is displayed on the chart.



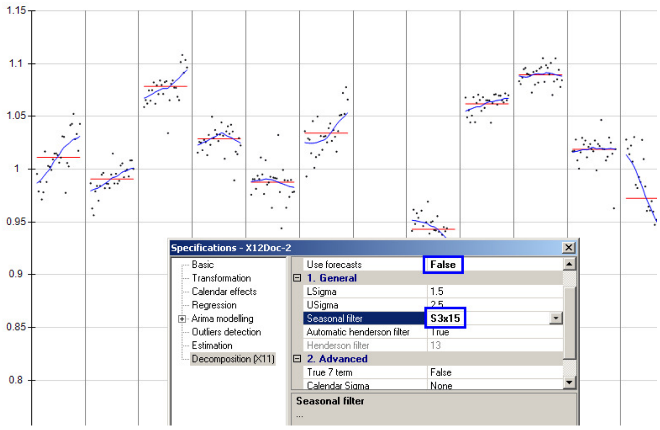

The trading days regression variables can be suppressed by setting the "Trading days -> Type" to "None" in the "Calendar effects" panel of the specification dialog box.




Meaningful information is provided in the "Pre-processing -> Arima" panel or in the different panels of the spectral analysis



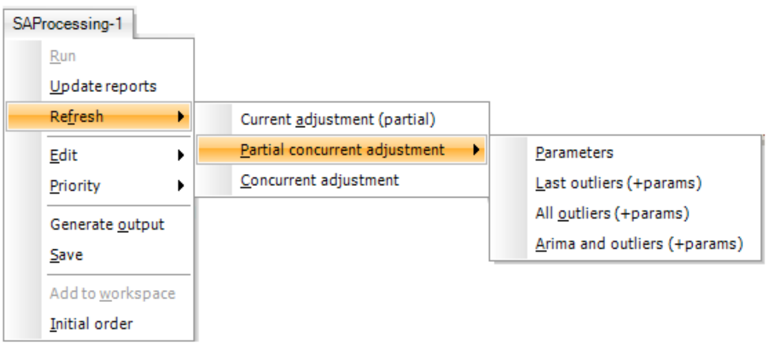

The X11 panel of the specification dialog box contains a rich set of options on the X11 decomposition. Their effects appear - for instance - in the SI-ratio chart.




The previous snapshot was realized by setting the "Use forecasts" option on false and the "Seasonal filter" on "S3x15"

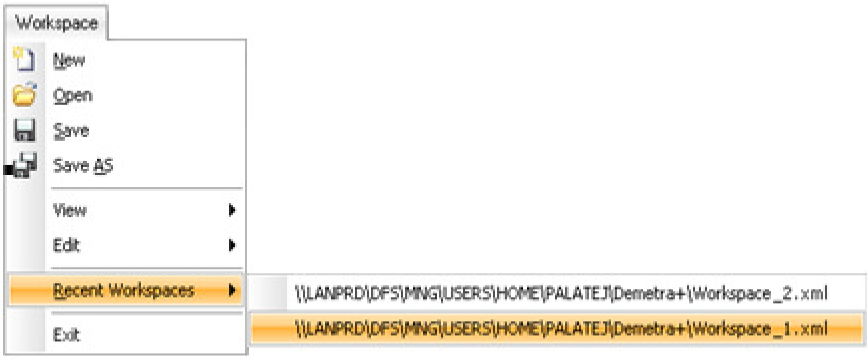
## 5.2 Saving and refreshing workspaces

By default, single and multi-processing generated through the so-called "short-ways" are not put in the current workspace. To be able to save and to refresh them, the user must first add them to the workspace. That can be done, for instance, through the main menu "SAProcessingXXX -> Add to Workspace".

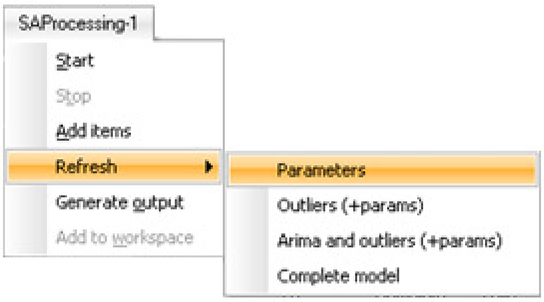


The user still has to save the workspace, using the usual menu command **Save**.

When Demetra+ is re-opened, it will automatically open at the last used workspace. The software also maintains a list of the most recently used workspace, which can be easily accessed.




A saved item of a workspace can be opened by a double click or by its local menu. It is then showed in its previous state. Demetra+ proposes several options to refresh it[28]:




| | |
|---|---|
| Parameters | Only the parameters are refreshed. The order of the ARIMA(p,d,q)(P,D,Q) is unchanged |
| Outliers (+ params) | Outliers and parameters are re-estimated |
| Last outliers (+params) | Outliers on the last periods and parameters are re-estimated |
| Complete | The model is completely re-estimated |

When the refresh option has been selected, Demetra+ automatically goes to the suitable time series provider(s) to ask for the updated observations; the new estimations are done on these series (using the previous models, modified by the chosen option).

28 For the moment, those options are only available for multi-processing.

# Annex

## 1. Definition of the residuals

Several tests are built on the residuals of the RegArima model. However, what we mean by "residuals" is not so obvious. TramoSeats and X12 use different definitions of residuals; Demetra+ proposes another one ("Stamp-like"). All those solutions correspond of course to the same likelihood (their sums of squares are identical) and they usually lead to very similar diagnostics. However, in some specific cases (short series, many regression variables and/or missing values...), they might lead to larger discrepancies.

We consider the ARIMA model:

$$\Phi(B)\Delta(B)(y_{orig} - X_{orig}\beta) = \mu + \Theta(B)\varepsilon$$

or, after differencing,

$$\Phi(B)(y - X\beta) = \mu + \Theta(B)\varepsilon$$

We consider that we have n observations after differencing and k regression variables.

Without going into all the mathematical details of the problem, below we shortly describe the different solutions[29].

### 1.1. TramoSeats

Tramo uses the following algorithm:

1. Filter the (differenced) exogenous variable by means of the Kalman filter. We get $y_l = L^{-1}y$, where $L$ corresponds to the lower triangular matrix of the Cholesky decomposition of the covariance matrix of the model.

2. Filter the (differenced) regression variables (including the additive outliers corresponding to missing values) by means of the Kalman filter. We get $X_l = L^{-1}X$

3. Solve by OLS the linear regression: $y_l = X_l\beta + \eta$

4. The OLS is solved by using the QR algorithm and the residuals are obtained through that procedure (see appendix 2 for further information on OLS estimation by QR decomposition).

[29] All estimations are based on the model after differencing, so that the noises are stationary ARMA processes.

The procedure provides n-k independent residuals. However, the relation to the time of those residuals is not obvious after the QR decomposition, so that the some tests (like periodograms) should not be used.

To get round that problem, Tramo also provides the "full residuals", defined by $y_l - X_l \hat{\beta}$ or, equivalently, by $L^{-1}(y - X\hat{\beta})$. The full residuals correspond to the one-step-ahead forecast error of the "linearized" series.

### 1.2. X12

X12 provides the exact maximum likelihood estimates (eml) of the residuals. However, that sentence has to be clarified.

We first consider the model without regression variables. When the stationary model is a pure MA model, it is easy to derive the maximum likelihood estimates of the residuals; they are defined for the period [-q, n[, where q is the order of the MA polynomial[30].

We give some more information on the handling of the general case, which is less documented. The software uses the following transformation:

$$z_t = \begin{cases} y_t, 0 \le t \le p \\ \Phi(B) y_t, t \ge p \end{cases}$$

where p is the auto-regressive order.

To simplify the notation, we will write below $y_O$ for $\{y_k\}_{0 \le k \le p}$ and $w$ for $\{z_k\}_{k \ge p}$.

We have that $p(y) = p(z) = p(y_O, w) = p(y_O | w) \cdot p(w)$.

$p(w)$ is solved by the solution developed for the pure MA case. It generates a set of $n - p + q$ residuals that correspond to the eml residuals of the transformed model. We write them $e_w$.

Using well-known properties of the normal distribution, we can derive the distribution of $p(y_O | w)$.Indeed, if , we have that

$$y \perp O | w \sim N(C^{'} \bullet [(\Omega \perp w)] \ (-1) w, [(\Omega)] \perp y - C^{'} \bullet [(\Omega \perp w)] (-1) \cdot C)$$

$C$ can be easily computed (using the Wald decomposition of the model) . The filtering algorithm of the pure MA model provides a transformation $A$ such that $A^{'} \cdot A = \Omega_w^{-1}$.

So, the initial residuals are computed as follows:

30 See LJUNG, G. and BOX, G. (1979) or OTTO, M. C., BELL, W. R., and BURMAN, J. P. (1987).

*1. Preliminary steps (independent of the observations)*

1.1 Compute $C$

1.2 Compute $A \cdot C$, by filtering the rows of $C$ with the MA algorithm

*2. Filtering*

2.1 Compute $w$

2.2 Compute $Aw$ with the ma algoritm

2.3 Multiply the results from 1.2 and 2.3 and subtract it from the ( $p$ ) first observations.

2.4 Pre-multiply those residuals by the inverse of the Cholesky decomposition of de variance matrix of $y \perp O \mid w$ (easily obtained from 1.2); the result correspond to the first $p$ "residuals". Those residuals can be interpreted as the one-step-ahead forecast error of $y_O$, knowing $w$. We note them $e_p$ .The complete set of the residuals are then $(e_p, e_w)$.

When the model contains regression variables, X12/X13 uses an iterative procedure: in a first step, for given coefficients, it computes the "linearized" series and it estimates by ml the parameters of the arima model for that series; in a second step, it re-estimates the coefficients of the regression parameters for the new model and it goes back to the first step. The final "residuals" are obtained from the linearized series by the procedure explained above when the iterative procedure has converged.

### 1.3. Demetra+

In Demetra+, the residuals are the one-step-ahead forecast errors of the state space model that contains the coefficients of the regression variables in the state vector (iterative gls model).

More precisely, we consider the following state space model:

$$y_t = Z_t \alpha_t$$

$$\alpha_t = (\tilde{\alpha}_t \ \beta_t)$$

$$Z_t = (\tilde{Z} \ X_t)$$

$$\alpha_{t+1} = \begin{pmatrix} \tilde{T} & 0 \\ 0 & 1 \end{pmatrix} \alpha_t + \begin{pmatrix} \tilde{R}\varepsilon_t \\ 0 \end{pmatrix}$$

where the tildes indicates the matrices of the ARIMA part.

In such a model, outliers have to be handled carefully: indeed, for each period corresponding to an outlier, the forecast error is missing (it cannot be estimated, the same way that initial residuals cannot be estimated when the model contains regression variables like calendar effects).

### 1.4. Final remarks

It should be noted that the original solution of Tramo and the solution which was implemented in Demetra+ are exactly equivalent when the model doesn't contain regression variables. The same is true for X12 only when the model is a pure AR model.

Finally, below we give a summary of the characteristics of the different solutions, considering the degrees of freedom of the residuals and their respect of the "time structure" (interpretation of the residuals in the time domain).

| Solution | Number of residuals | Independence of the residuals | Respect of the time structure |
|---|---|---|---|
| Tramo ("QR residuals") | $n-k$ | x | |
| Tramo ("full residuals") | $n$ | | x |
| X12 | $n+q$ | | x (partially) |
| Demetra+ | $n-k$ | x | x |

## 2. Least squares estimation by means of the QR decomposition.

We consider the regression model

$$y = X\beta + \varepsilon$$

The least squares problem consists in minimizing the quantity

$$\| X\beta - y \|_2^2$$

Provided that the regression variables are independent, it is possible to find an orthogonal matrix $Q$, so that

$Q \cdot X = \begin{pmatrix} R \\ 0 \end{pmatrix}$ where $R$ is upper triangular.

That matrix is built by means of Householder transformations (reflections).

We have now to minimize

$$\left\|QX\beta = Qy\right\|_2^2 = \left\| \begin{pmatrix} R \\ 0 \end{pmatrix} \beta - Qy \right\|_2^2 = \| R\beta - a \|_2^2 + \| b \|_2^2$$

where

$(Qy)_{0...x-1} = a$ and $(Qy)_{x...n-1} = b$ .

It is trivially done by setting

$\beta = R^{-1}a$ . In that case, $\| R\beta - a \|_2^2 = 0$

The "residuals" obtained by that procedure are then $b$ , as defined above.

It should be noted that the QR factorization is not unique, and that the final residuals also depend on the order of the regression variables (the columns of $X$ ).

## 3. *Specifications*

| SA Method | Name | Settings |
|---|---|---|
| TramoSeats | RSA0 | Level, Airline model |
| | RSA1 | Log/level, outliers detection, Airline model |
| | RSA2 | Log/level, working days, Easter, outliers detection, Airline model |
| | RSA3 | Log/level, outliers detection, automatic model identification |
| | RSA4 | Log/level, working days, Easter, outliers detection, automatic model identification |
| | RSA5 | Log/level, trading days, Easter, outliers detection, automatic model identification |
| X12 | X11 | No pre-processing |
| | RSA1 | Log/level, outliers detection, Airline model |
| | RSA2c | Log/level, working days, Easter, outliers detection, Airline model, pre-adjustment for leap-year if logarithmic transformation has been used |
| | RSA3 | Log/level, outliers detection, automatic model identification |
| | RSA4c | Log/level, working days, Easter, outliers detection, automatic model identification, pre-adjustment for leap-year if logarithmic transformation has been used |
| | RSA5 | Log/level, trading days, Easter, outliers detection, automatic model identification, pre-adjustment for leap-year if logarithmic transformation has been used |

***Explanations for settings:***

Level – no transformation is performed

Log/level – Demetra+ tests for the log/level specification,

Working days – a pretest is made for the presence of Working Day using one parameter specification (working vs. non-working days),

Trading days – a pretest is made for the presence of Trading Day using six parameters specification (for working days, the day of week: Monday,...,Fridayis specified),

Easter – the program tests for the necessity of a correction for Easter effect in the original series.

Outliers detection – Demetra+ automatically detects all types of outliers including: AO (additive outliers), LS (level shifts), TC (transitory outliers) using default critical values.

Airline model – an Airline model (0,1,1)(0,1,1) is estimated.

Automatic model identification – Demetra+ identifies and estimates the best ARIMA model.

## 4. Tests

### 4.1. Doornik-Hansen test

The Doornik-Hansen is defined as follows:

let s = skweness, k=kurtosis of the n (non missing) residuals

We make the following transformations:

*Transformation of the skewness (D'Agostino)*

$$\beta = \frac{3(n^2+27n-70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)}$$

$$\omega^2 = -1 + \sqrt{2(\beta - 1)}$$

$$\delta = \frac{1}{\sqrt{0.5 \log \omega^2}}$$

$$y = s\sqrt{\frac{(\omega^2-1)(n+1)(n+3)}{12(n-2)}}$$

$$z_1 = \delta \log(y + \sqrt{y^2 - 1})$$

*Transformation of the kurtosis (Wilson-Hilferty)*

$$\delta = (n-3)(n+1)(n^2+15n-4)$$

$$a = \frac{(n-2)(n+5)(n+7)(n^2+27n-70)}{6\delta}$$

$$c = \frac{(n-7)(n+5)(n+7)(n^2+2n-5)}{6\delta}$$

$$l = \frac{(n+5)(n+7)(n+7)(n^3+37n^2+11n-313)}{12\delta}$$

$$\alpha = a + c \cdot s \cdot s$$

$$\chi = 2l(k-1-s^2)$$

$$z_2 = (\sqrt{9\alpha})\left(\frac{1}{9\alpha} - 1 + \sqrt[3]{\frac{\chi}{2\alpha}}\right)$$

$$DH = z_1^2 + z_2^2 \sim \chi^2(2)$$

### 4.2. Ljung-Box test

The Ljung-Box test is defined as follows:

let $\rho_j$ the sample autocorrelation at rank j of the n residuals. The Ljung-Box statistics is

$$LB(k) = n \cdot (n-2)\sum_{j=1}^{k} \frac{\rho_j^2}{n-j}$$

If the residuals are random, it should be distributed as $\chi^2(k-np)$ where $np$ is the number of hyper-parameters of the model from which the residuals are derived.

### 4.3. Spectral test

#### 4.3.1. Definition of the periodogram

The periodogram of the series $\{y_t\}_{1<t\leq n}$ is computed as follows:

1. The $y_t$ is standardized

$$\bar{y} = \frac{\sum_{t=1}^{t\leq n} y_t}{n}$$

$$\hat{\delta}_y^2 = \frac{\sum_{t=1}^{t\leq n}(y_t - \bar{y})^2}{n}$$

$$z_t = \frac{(y_t - \bar{y})}{\hat{\delta}_y}$$

2. The periodogram is computed on the standardized $z_t$

$$I_{n,z}(\lambda) = \frac{2}{n}(C^2_{n,z}(\lambda) + S^2_{n,z}(\lambda))$$

where

$C_{n,z}(\lambda) = \sum_{t=1}^{n} \cos(\lambda t) z_t$ and $S_{n,z}(\lambda) = \sum_{t=1}^{n} \sin(\lambda t) z_t$

### 4.3.2. Periodogram at the Fourier frequencies

The Fourier frequencies are defined by

$$\lambda_j = \frac{2\pi j}{n}, 0 < j \leq \left[\frac{n}{2}\right]$$

If the $z_t$ are iid $N(0,1)$, it is easy to see that the corresponding quantities $I_{n,z}(\lambda_j)$ are iid $\chi^2(2)$.

We have indeed that

$$\sum_{t=1}^{n} e^{it(\lambda_j - \lambda_k)} = \begin{cases} n \text{ if } j = k \\ 0 \text{ if } j \neq k \end{cases}$$

and

$$\sum_{t=1}^{n} \cos^2(\lambda_j t) = \sum_{t=1}^{n} \sin^2(\lambda_j t) = \frac{n}{2},,$$

so that $\sqrt{\frac{2}{n}} C_{n,z}(\lambda_j)$ and $\sqrt{\frac{2}{n}} S_{n,z}(\lambda_k)$ are uncorrelated $N(0,1)$ random variables.

### 4.3.3. Test on the periodogram

Under the hypothesis that $z_t$ is a Gaussian white noise, and considering subset J of Fourier frequencies, we have:

$$\Pr\left\{\max_{j\in J} I_{n,z}(\lambda_j) \le \alpha\right\} = \left[1 - e^{-\frac{\alpha}{2}}\right]^{\#J}$$

If we consider the sets of Fourier frequencies on or near the trading days frequencies on one side and on or near the seasonal frequencies on the other side, we can use the above formula as rough test regarding the absence of trading days/seasonal effects in the considered series.

The software considers the Fourier frequencies which are on or near the following frequencies (the nearest is chosen, or two if they are equidistant):

| Annual frequency | Seasonal | Trading days |
|---|---|---|
| 12 | 2π/12, 4π/12, 6π/12, 8π/12, 12π/12 | d, 2.714 |
| 6 | 2π/6, 4π/6 | d |
| 4 | 2π/4 | d, 1.292, 1.850, 2.128 |
| 3 | - | d |
| 2 | - | d |

where d is computed as follows:

if s is the frequency of the series,

$$n = \frac{365.25}{s}$$

$$d = \frac{2\pi}{7} \cdot \textit{(n modulo 7)}$$

### 4.4. Seasonality tests

This section presents the set of seasonality tests calculated by Demetra+. Detailed description of these tests and testing procedure is available in LADIRAY D. and QUENNEVILLE B. (1999).

#### 4.4.1. Friedman test (stable seasonality test)

Friedman's test is a non-parametric method for testing that samples are drawn from the same population or from populations with equal medians. In the regression equation the significance of the month (or quarter) effect is tested. Friedman test requires no distributional assumptions. It uses the rankings of the observations.

Seasonal adjustment procedures uses Friedman test for checking the presence of seasonality. Friedman test is called a stable seasonality test. This test uses preliminary estimation of the unmodified Seasonal-Irregular component[31] (for X12 this time series is shown in table B3) from which $k$ samples are derived ($k = 12$ for monthly series and $k = 4$ for quarterly series) of size $n_1, n_2, \ldots n_k$ respectively. Each $k$ corresponds to a different level of seasonality. It is assumed that seasonality affect only the means of the distribution and not their variance. Assuming that each

31 Unmodified Seasonal-Irregular component is the seasonal-irregular factors with the extreme values.

sample is derived from a random variable $X_j$ following the normal distribution with mean $m_j$ and standard deviation $\sigma$ the following null hypothesis is tested:

$H_0: \quad m_1 = m_2 = ... = m_k$

against:

$H_1: \quad m_p \neq m_q$ for the least one pair $(p, q)$

The test uses the following decomposition of the variance:

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{i,j} - \bar{x})^2 = \sum_{j=1}^{k} n_j(\bar{x}_{\cdot j} - \bar{x}_{\bullet\bullet})^2 + \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{i,j} - \bar{x}_{\bullet j})^2$$

where:

$\bar{x}_k$ - the average of $k$ -th sample.

The total variance is therefore broken down into a variance of the averages due to seasonality and a residual seasonality.

The test statistics is calculated as:

$$F_S = \frac{\dfrac{\sum_{j=1}^{k} n_j(\bar{x}_{\bullet j} - \bar{x}_{\bullet\bullet})^2}{k-1}}{\dfrac{\sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{i,j} - \bar{x}_{\bullet j})^2}{n-k}} \sim F(k-1, n-k)$$

Where $k-1$ and $n-k$ are degrees of freedom.

The number of observations in preliminary estimation of the unmodified Seasonal-Irregular is lower than in final estimation of the unmodified Seasonal-Irregular component. Because of that the number of degrees of freedom in stable seasonality test is lower than number of degrees of freedom in test for the presence of seasonality assuming stability (see 4.4.3) (e.g. X12 uses centered moving average of order 12 to calculate the preliminary estimation of trend-cycle. As a result the first six and last six points in the series are not computed at this stage of calculation. Preliminary estimation of trend-cycle is then used for calculation the preliminary estimation of the unmodified Seasonal-Irregular).

If the null hypothesis of no stable seasonality is not rejected at the 0.10% significance level ($P_s \geq 0.001$), then the series is considered to be non-seasonal.

### 4.4.2. Kruskal-Wallis test

Kruscal-Wallis test is a non-parametric test used for comparing samples from two or more groups. The null hypothesis states that all months (or quarters, respectively) have the same mean.

The test is calculated for the final estimation of the unmodified Seasonal-Irregular component from which $k$ samples are derived ($k = 12$ for monthly series and $k = 4$ for quarterly series) of size $n_1, n_2, \ldots n_k$ respectively.

The test is based on the statistic:

$$W = \frac{12}{n(n+1)} \sum_{j=1}^{k} \frac{S_j^2}{n_j} - 3(n+1)$$

Where $S_j$ is the sum of the ranks of the observations from the sample $A_j$ within the whole sample of $n = \sum_{j=1}^{k} n_j$ observations.

Under the null hypothesis the test statistic follows a chi-square distribution with $k-1$ degrees of freedom.

### 4.4.3. Test for the presence of seasonality assuming stability

The test statistics and testing hypothesis are the same as for Friedman stable seasonality test described in 4.4.1 (Annex). The test statistics is calculated for final estimation of the unmodified Seasonal-Irregular Component (in case of X12 this series is presented in table D8).

### 4.4.4. Evaluative seasonality test (Moving seasonality test)

The test is based on a two-way analysis of variance model. The model uses the values from complete years only. For the seasonal-irregular component it uses one of the following models, depending on the decomposition's type:

Multiplicative:

$$\left| SI_{ij} - 1 \right| = X_{ij} = b_i + m_j + e_{ij}$$

Additive:

$$\left| SI_{ij} \right| = X_{ij} = b_i + m_j + e_{ij}$$

Where $m_j$ refers to the monthly or quarterly effect for $j$-th period,

$b_j$ refers to the annual effect $i$ ($i = 1, \ldots, N$) where $N$ is the number of complete years,

$e_{ij}$ represents the residual effect.

The test is based on the decomposition $S^2 = S_A^2 + S_B^2 + S_R^2$ where:

$S_R^2 = \sum_{i=1}^{N}\sum_{i=1}^{N}(\bar{X}_{ij} - \bar{X}_{\bullet\bullet})^2$ - the total sum of squares,

$S_A^2 = k\sum_{i=1}^{N}(\bar{X}_{\bullet j} - \bar{X}_{\bullet\bullet})^2$ - the inter-month (inter-quarter, respectively) sum of squares,

$S_B^2 = k\sum_{i=1}^{N}(\bar{X}_{i\bullet} - \bar{X}_{\bullet\bullet})^2$ - the inter-year sum of squares,

$S_R^2 = \sum_{i=1}^{N}\sum_{i=1}^{N}(\bar{X}_{ij} - \bar{X}_{i\bullet} - \bar{X}_{\bullet j} + \bar{X}_{\bullet\bullet})^2$ - the residual sum of squares.

The null hypothesis $H_0$ is that $b_1 = b_2 = ... = b_N$ which means that there is no change in seasonality over the years.

This hypothesis is verified by the following test statistics:

$$F_M = \frac{\frac{S_B^2}{(N-1)}}{\frac{S_R^2}{(N-1)(k-1)}}$$

which follows a $F$ -distribution with $k-1$ and $n-k$ degrees of freedom.

### 4.4.5. Test for presence of identifiable seasonality

This test combines the F-statistic values of parametric test for stable seasonality and for the moving seasonality described above.

The test statistic is:

$$T = \left(\frac{\frac{7}{F_S} + \frac{3F_M}{F_S}}{2}\right)^{0.5}$$

Where $F_S$ is a stable seasonality test statistic and $F_M$ is moving seasonality test statistic.

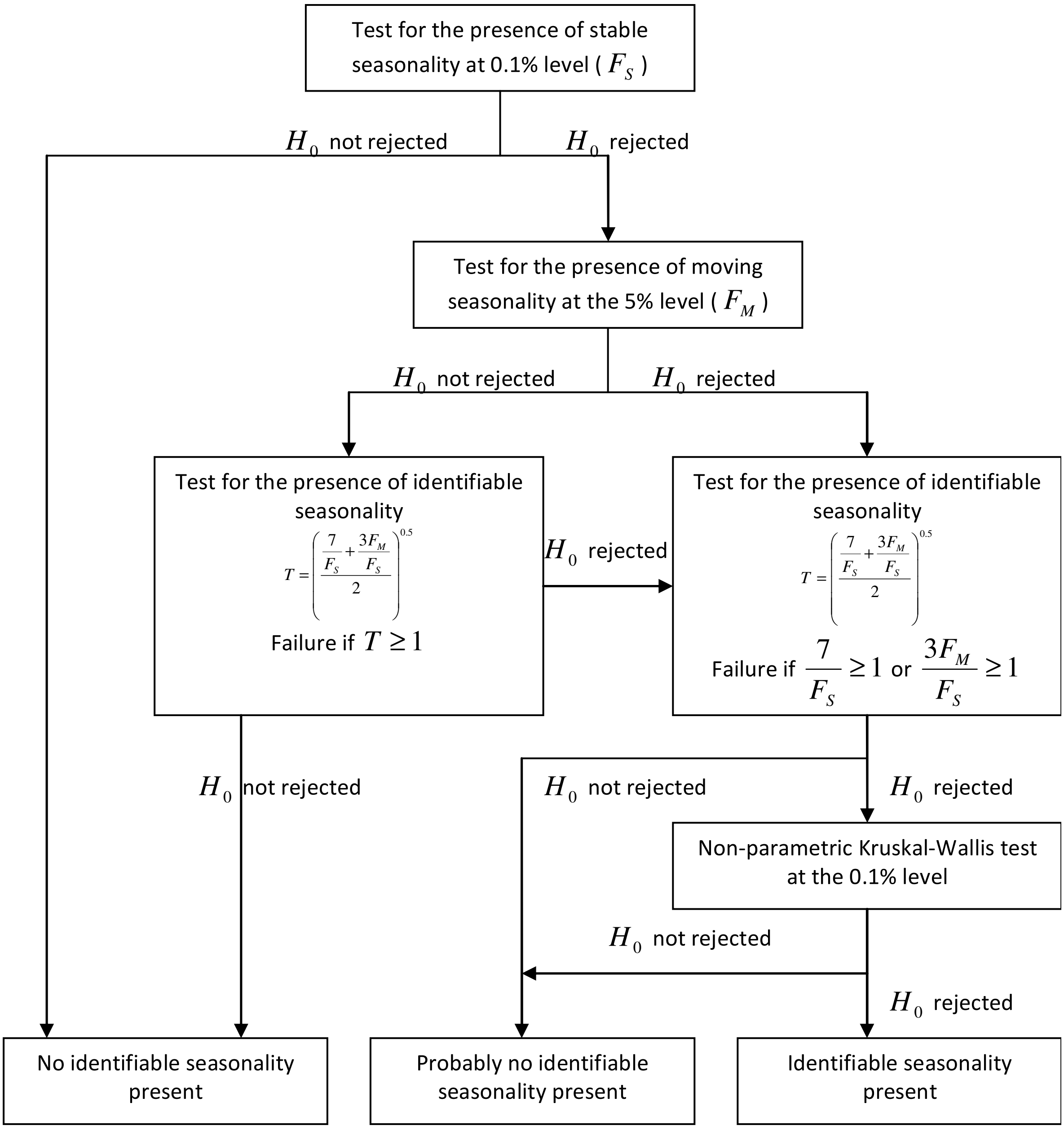
The detailed test's description is available in LOTHIAN J., and MORRY M. (1978).

### 4.4.6. Combined seasonality test

This test combines the Kruskal-Wallis test (4.4.2) along with test for the presence of seasonality assuming stability (4.4.3), evaluative seasonality test (4.4.4) and test for presence of identifiable seasonality (4.4.5). All those test are calculated using final unmodified SI component. The main purpose of combined seasonality test is to check whether the seasonality of the series is

identifiable. For example, identification of the seasonal pattern is problematic if the process is dominated by highly moving seasonalit[32]y.

The testing procedure is shown below:




[32] DAGUM, E. B. (1987).

## *5. X-12-ARIMA – tables*

**Part A - Preliminary Estimation of Extreme Values and Calendar Effects**

Table A1- Original series

Table A1a – Forecast of Original Series

Table A2 – Leap year effect

Table A6 – Trading Day effect (1 or 6 variables)

Table A7 – Easter effect

Table A8 – Total Outlier Effect

Table A8ao – Additive outlier effect

Table A8ls – Level shift effect

Table A8tc – Transitory effect

**Part B - Preliminary Estimation of Time Series Components**

Table B1 - Original series

Table B2 - Unmodified Trend-Cycle

Table B3 - Unmodified Seasonal-Irregular Component

Table B4 - Replacement Values for Extreme SI Values

Table B5 - Seasonal Component

Table B6 - Seasonally Adjusted Series

Table B7 - Trend-Cycle

Table B8 - Unmodified Seasonal-Irregular Component

Table B9 - Replacement Values for Extreme SI Values

Table B10 - Seasonal Component

Table B11 - Seasonally Adjusted Series

Table B13 - Irregular Component

Table B17 - Preliminary Weights for the Irregular

Table B20 - Adjustment Values for Extreme Irregulars

**Part C - Final Estimation of Extreme Values And Calendar Effects**

Table C1 - Modified Raw Series

Table C2 -Trend-Cycle

Table C4 - Modified SI

Table C5 - Seasonal Component

Table C6 - Seasonally Adjusted Series

Table C7 - Trend-Cycle

Table C9 - SI Component

Table C10 - Seasonal Component

Table C11 - Seasonally Adjusted Series

Table C13 - Irregular Component

Table C20 - Adjustment Values for Extreme Irregulars

**Part D - Final Estimation of the Different Components**

Table D1 - Modified Raw Series

Table D2 -Trend-Cycle

Table D4 - Modified SI

Table D5 - Seasonal Component

Table D6 - Seasonally Adjusted Series

Table D7 - Trend-Cycle

Table D8 - Unmodified SI Component

Table D9 - Replacement Values for Extreme SI Values

Table D10 - Final Seasonal Factors

Table D10A – Forecast of Final Seasonal Factors

Table D11 - Final Seasonally Adjusted Series

Table D11A - Final Seasonally Adjusted Series with Revised Annual Totals

Table D12 - Final Trend-Cycle

Table D12A – Forecast of Final Trend Component

Table D13 - Final Irregular Component

Table D13U – Irregular component (excluded outlier effects)

Table D16 - Seasonal and Calendar Effects

Table D16A – Forecast of Seasonal and Calendar Component

Table D18 - Combined Calendar Effects Factors

**Part E - Components Modified for Large Extreme Values**

Table E1 - Raw Series Modified for Large Extreme Values

Table D2 - SA Series Modified for Large Extreme Values

Table E3 - Final Irregular Component Adjusted for Large Extreme Values

Table E11 - Robust Estimation of the Final SA Series

## 6. *Visual spectral analysis*

The autoregressive spectrum estimator is defined as follows[33]:

$$\hat{s}(\lambda) = 10\log_{10}\left\{\frac{\hat{\delta}_m^2}{2\pi\left|1-\sum_{j=1}^{m}\hat{\phi}_j e^{i2\pi j\lambda}\right|^2}\right\}e^{i2\pi j\lambda}$$

where:

33 Definition taken from: 'X-12-ARIMA Reference Manual', p. 55, http://www.census.gov/srd/www/x12a/.

$\lambda$ - frequency, $0 \le \lambda \le 0.5$,

$\hat{\delta}^2{}_m$ - the sample variance of the residuals,

$\hat{\phi}_j$ - coefficients from regression $x_t - \bar{x}$ on $x_{t-j} - \bar{x}$, $1 \le j \le m$.

Criterion of "visual significance" is based on the range $\hat{s}^{\max} - \hat{s}^{\min}$ of the $\hat{s}(\lambda)$ values,

where:

$\hat{s}^{\max} = \max_k \hat{s}(\lambda_k)$,

$\hat{s}^{\min} = \min_k \hat{s}(\lambda_k)$.

The value is considered to be visually significant if $\hat{s}(\lambda_k)$ at a trading day or seasonal frequency $\lambda_k$ (other than the seasonal frequency $\lambda_{60}$ = 0.5) must be above the median of the plotted values of $\hat{s}(\lambda_k)$ and must be larger than both neighboring values $\hat{s}(\lambda_{k-1})$ and $\hat{s}(\lambda_{k+1})$ by at least 6/52 times the range $\hat{s}^{\max} - \hat{s}^{\min}$ .

For a given series $y_{t, 0 \le t < T}$, which may contain missing values, the periodogram is computed as follows:
In a first step, the series is standardized:

$$z_t = \frac{y_t - \hat{y}_t}{\sigma(y_t)}$$

In a second step, we compute at the so-called Fourier frequencies $(\omega = \frac{2\pi i}{T}, 0 \le i < \frac{T+1}{2})$ which are the values of the periodogram:

$$\frac{2}{N} \cdot \left| \sum_{t=0, z_t \text{ defined}}^{t<T} z_t e^{i\omega t} \right|$$

where $N$ is the number of non missing values.
Under the white noise hypothesis, the values of the periodogram should be asymptotically distributed as a Chi-square with 2 degrees of freedom.
The default frequency $td$ for trading days is computed as follows (for series of quarterly series):

$$n = \frac{365{,}25}{q}, q = 4$$

$$td = \frac{2\pi}{7}\left(n - 7 \cdot \left[\frac{n}{7}\right]\right)$$

Other frequencies correspond to trading days frequencies:

- For monthly series, 2.714 ( default = 2.188)
- For quarterly series, 1.292, 1.850, 2.128 (default = 0.280)

## 7. Revision histories

Revisions are calculated as a difference between the first (earliest) adjustment of an observation computed when that observation is the final period of the time series (*concurrent adjustment,* denotes as $A_{t|t}$ ) and a later adjustment based on all data span (*most recent adjustment,* denotes as $A_{t|N}$ ).

In case of multiplicative decomposition the revision history of the seasonal adjustment from time $N_0$ to $N_1$ is a sequence of $R^A_{t|N}$ calculated in a following way[34]:

$$R^A_{t|N} = 100 \times \frac{A_{t|N} - A_{t|t}}{A_{t|t}}$$

The revision history of the trend is calculated in a similar way:

$$R^T_{t|N} = 100 \times \frac{T_{t|N} - T_{t|t}}{T_{t|t}}$$

With additive decomposition $R^A_{t|N}$ is calculated in the same way if all values $A_{t|t}$ have the same sign[35]. Otherwise differences are calculated as:

$$R^A_{t|N} = A_{t|N} - A_{t|t}$$

The analogous quantities are calculated for final Henderson trends.

## 8. Sliding spans

Each period (month or quarter) which belongs to more than one span is examined to see if its seasonal adjustments vary more than a specified amount across the spans.

Seasonal factor is regarded to be unreliable if the following condition is fulfilled:

34 FINDLEY, D. F., MONSELL, B. C., BELL, W. R., OTTO, M. C., and CHEN, B-C. (1998).

35 'X-12-Arima Reference Manual' (2007).

$$SS_t = \frac{\max_{k \in N_t} S_t(k) - \min_{k \in N_t} S_t(k)}{\min_{k \in N_t} S_t(k)} > 0.03,$$

Where:

$S_t(k)$ - the seasonal factor estimated from span $k$ for month $t$.

$N_t$ = {$k$ : period $t$ is in the $k$-th span}.

For seasonally and trading days adjusted series the following statistic is being calculated:

$$\frac{\max_j A_t^j - \min_j A_t^j}{\min_j A_t^j}$$

The value is considered to be unreliable if it is higher than $0.03$.

Similarly, the seasonally adjusted changes are unstable if:

$$\max_j \frac{A_t^j}{A_{t-1}^j} - \min_j \frac{A_t^j}{A_{t-1}^j} > 0.03$$

Where:

$A_t(k)$ - the seasonally (or trading day) adjusted value from span $k$ for month $t$

$N1(t)$ = {$k$ : period $t$ and $t-1$ are in the $k$-th span}.

## 9. *Code to generate simple seasonal adjustments (C#)*

(Some namespaces have been removed to simplify the reading)

```
// creates a new time series
// parameters: frequency/first year/first period (0-based)/array of doubles/copy
data (uses the current array or creates a copy)
TSData s = new TSData(12, 1967, 0, g_prodind, false);

// basic processing

// tramo-seats specification. RSA5 (full automatic)
TramoSeats.Specification ts_spec =     TramoSeats.Specification.RSA5;

// launches tramo-seats core engine
TramoSeats.Monitor ts_monitor=new TramoSeats.Monitor();
// executes the processing
TramoSeats.TramoSeatsResults ts_rslts = ts_monitor.Process(s, ts_spec);

// x12 specification. equivalent RSA5 (full automatic)
X12.Specification x_spec = X12.Specification.RSA5;

// launches tramo-seats core engine
X12.Monitor x_monitor=new X12.Monitor();
// executes the processing
X12.X12Results x_rslts = x_monitor.Process(s, x_spec);
```

```
// seasonally adjusted series
TSData ts_sa = ts_rslts.Series(SAComponentType.CSA);
TSData x_sa = x_rslts.Series(SAComponentType.CSA);

// computes diffrences between both results...
TSData diff = ts_sa - x_sa;

// computes statistics on the differences...
DescriptiveStatistics stats = new DescriptiveStatistics(diff.Values);

double max = stats.Max, min = stats.Min, rmse = Math.Sqrt(stats.SumSquare /
diff.Length);

// more advanced uses (computed "on the fly")
Periodogram periodogram = new Periodogram(x_rslts.X11Results.DTables["D8"]
       .Values);

// roots of the moving average polynomial of the arima model used by Seats
Complex[] roots = ts_rslts.Seats.SArima.MA.Roots();
```

# REFERENCES


BOX, G. E. P. and TIAO, G. C. (1975), 'Intervention Analysis with Applications to Economic and Environmental Problems', *Journal of the American Statistical Association*, No. 70.

CAPORELLO, G., and MARAVALL, A. (2004), 'Program TSW. Revised Reference Manual', Banco de España. http://www.bde.es/servicio/software/tramo/tswrm.pdf).

DAGUM, E. B. (1987), 'Modelling, Forecasting and Seasonally Adjusting Economic Time Series with the X-11 ARIMA Method', *Journal of the Royal Statistical Society*, Series D (The Statistician), Vol. 27, No. 3/4.

DAGUM, E. B. (1979), 'On the Seasonal Adjustment of economic Time Series Aggregates: A Case Study of the Unemployment Rate', *Counting the Labor Force, National Commission Employment and Unemployment Statistics, Appendix*, 2, 317-344, Washington.

DURBIN, J., and KOOPMAN, S. J. (2001), 'Time Series Analysis by State Space Methods'. Oxford University Press.

(2000), 'Seasonal Adjustment of Monetary Aggregates and HICP for the Euro Area', ECB, August 2000, http://www.ecb.int/pub/pdf/other/sama0008en.pdf.

FINDLEY, D., and MARTIN, D. (2006), 'Frequency Domain Analyses of SEATS and X–11/12-ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series', *Journal of Official Statistics*, Vol.22, No.1, 2006. pp. 1–34

FINDLEY, D. F., MONSELL, B. C., BELL, W. R., OTTO, M. C., and CHEN, B-C. (1998), 'New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program', *Journal of Business and Economic Statistics*, 2, 16, 127-152.

FINDLEY, D., MONSELL, B. C., SHULMAN, H. B., and PUGH, M. G. (1990), 'Slidings Spans Diagnostics for Seasonal and Related Adjustments', *Journal of the American Statistical Association*, vol. 85, n° 410.

GOMEZ, V., and MARAVALL, A. (1994), 'Estimation, Prediction, and Interpolation for Nonstationary Series with the Kalman Filter', *Journal of the American Statistical Association*, vol. 89, n° 426, 611-624.

HARVEY, A. (1989), '*Forecasting, Structural Time Series Models and the Kalman Filter*', Cambridge University Press.

GÓMEZ, V., and MARAVALL, A. (2001), '*Seasonal Adjustment and Signal Extraction in Economic Time Series*', in '*A Course in Advanced Time Series Analysis*', PEÑA, D., TIAO, G., and TSAY, R. (eds.), Wiley and Sons, New York, 202-246.

(2007), 'Guide to Seasonal Adjustment', ONS Methodology and Statistical Development.

HYLLEBERG, S., ed., (1992), 'Modelling Seasonality'. Oxford, New York, Toronto, Oxford University Press.

KAISER, R., and MARAVALL, A. (1999), 'Seasonal Outliers in Time Series'. Documento de Trabajo 9915, Banco de España.

KAISER, R. and MARAVALL, A. (2000), 'Notes on Time series Analysis, ARIMA Models and Signal Extraction', Banco de España Working Papers, No 12, Banco de España.

LADIRAY, D., and QUENNEVILLE B. (1999), 'Seasonal Adjustment with the X-11 Method', Lecture Notes in Statistics, 2001, New York: Springer-Verlag.

LADIRAY, D. and MAZZI, G. L. (2003), 'Seasonal Adjustment of European Aggregates: Direct versus Indirect Approach', Proceedings of the Seminar on Seasonal Adjustment, MANNA, M., and PERONACI, R. ed., European Central Bank, 37-66.

LOTHIAN, J., and MORRY, M. (1978), 'A Test for Presence of Identifiable Seasonality when Using the X-11 Program', Working Paper, Time Series Research Division, Statistics Canada, Ottawa, ON, Canada.

LJUNG, G., and BOX, G. (1979), 'The likelihood function of stationary autoregressive-moving average models', *Biometrika* 66, 265–270.

MARAVALL, A. (1993), 'Stochastic linear trends', *Journal of Econometrics*, 56 (1993) 5-37.

MARAVALL, A. (2008), 'Notes On Programs TRAMO and SEATS: TRAMO part' http://www.bde.es/webbde/es/secciones/servicio/software/tramo/Part_II_Tramo.pdf

MARAVALL, A. (2006), 'An application of the TRAMO-SEATS automatic procedure; direct versus indirect adjustment', *Computational Statistics & Data Analysis* 50 (2006) 2167 – 2190.

MARAVALL, A. (1995), 'Unobserved Components in Economic Time Series', in H. Pesaran and M. Wickens (eds.), The Handbook of Applied Econometrics, Oxford.

OTTO, M. C., BELL, W. R., and BURMAN J. P. (1987), 'An iterative GLS approach to maximum likelihood estimation of regression models with ARIMA errors', Research Report No. 87/34, Statistical Research Division, Bureau of the Census.

SHISHKIN, J., YOUNG, A. H., and MUSGRAVE, J. C. (1967), 'The X-11 variant of the Census Method II seasonal adjustment program', Technical Paper, No. 15, U.S. Bureau of the Census.

SOKUP, R. J., and FINDLEY, D. F. (1999), 'On the Spectrum Diagnostics Used by X-12-Arima to Indicate the Presence of Trading Day Effects after Modeling or Adjustment', *Proceedings of the American Statistical Association*, Business and Economic Statistics Section.

(2006), 'Seasonal Adjustment Filters for Short and Moderate-Length Time Series', *Journal of Official Statistics*, Vol. 22, No. 1, http://www.census.gov/ts/papers/findleymartinjosreprint.pdf.

(2007), 'X-12-ARIMA Reference Manual. Time Series', Staff Statistical Research Division, U.S. US Bureau of the Census, (http://www.census.gov/srd/www/x12a/x12down_pc.html).